

[illegible][illegible]

1950 1960 1970 1980 1990 2000 2010
CAGAAATGCAAGATGAAGTCAAAACTGTTGGTGAGGAACAAAAGCCAGAGAAACCAAGCGTAGGGTGGGATA
1920 1930 1940 1950 1960 1970 1980
CAGAAATGCAAGATGAAGTCAAAACTGTTGGTGAGGAACAAAAGCCAGAGAAACCAAGCGTAGGGTGGGATA
1990 2000 2010 2020 2030 2040 2050
CTTCAGTGTCTTGGGAAGCACTGATTTGTTCGGATCATCCAAAGAGAGAGCAAGGAAGCATCTCTTCAG
2020 2030 2040 2050 2060 2070 2080
CTTCAGTGTCTTGGGAAGCACTGATTTGTTCGGATCATCCAAAGAGAGAGCAAGGAAGCATCTCTTCAG
2090 2100 2110 2120 2130 2140 2150 2160
ATGATCAAGAGAGGCCCAAGACACTGGAGGGGACAGTCAACAGAGAGAGAGGCGCAGCAAGAAAGAAAG
2170 2180 2190 2200 2210 2220 2230
CGGAAACAGACGCTGTTCTTCGCCAGCACCCAGAGAGAGGACCAAGCGCAAGGAAGTTCTCTCACCCGAGCCAG
2130 2140 2150 2160 2170 2180 2190 2200
CGGAAACAGACGCTGTTCTTCGCCAGCACCCAGAGAGAGGACCAAGCGCAAGGAAGTTCTCTCACCCGAGCCAG
2210 2220 2230 2240 2250 2260 2270
CGGAAAGCCCTTCCGAAGGGGAAGGTGTTCCACTTGGGAGTCAATTTAAAGATTAGTCACTCCAAAGAAAA
2280 2290 2300 2310 2320 2330 2340
AATCCAAAGTCAAAACTGGAAGAGAAAGCCGAAGACTCTAGTGTAGAGCAGTTGTCTCCACTGAGATCGAACCGA
2310 2320 2330 2340 2350 2360 2370
AATCCAAAGTCAAAACTGGAAGAGAAAGCCGAAGACTCTAGTGTAGAGCAGTTGTCTCCACTGAGATCGAACCGA
2380 2390 2400 2410 2420 2430 2440
GTAGAGAAGAACTCTGGGTTTCATTTAAGAAATTCATCCCGGACCGCGGAAGAAAGGCGACAGCGGAAGC
2450 2460 2470 2480 2490 2500 2510 2520
AAGAAACAGCCACTGTGGAAGACTCAGGGCCAGTGGAGATAAATGAGGACGACCTTAATGTCTCCAGCCGCTCG
2420 2430 2440 2450 2460 2470 2480
AAGAAACAGCCACTGTGGAAGACTCAGGGCCAGTGGAGATAAATGAGGACGACCTTAATGTCTCCAGCCGCTCG
2530 2540 2550 2560 2570 2580 2590
TGCCCTGTCTGAGTATAATGCAAGTGGAGGGGAGAGATGGAAGCCCAAGGGGAATACCGAGGTGCCCCAGC
2490 2500 2510 2520 2530 2540 2550 2560
TGCCCTGTCTGAGTATAATGCAAGTGGAGGGGAGAGATGGAAGCCCAAGGGGAATACCGAGGTGCCCCAGC
2600 2610 2620 2630 2640 2650 2660
TGCTGGGGCTGTGACGTGTCGAGGAGCTCAGTAAGACTCTGTCACACTGTGAGTGTCCGACTCATTTG
2570 2580 2590 2600 2610 2620 2630
TGCTGGGGCTGTGACGTGTCGAGGAGCTCAGTAAGACTCTGTCACACTGTGAGTGTCCGACTCATTTG
2670 2680 2690 2700 2710 2720 2730
ATGGGACACAGGGCAGTCAACAGTGTGGAAGGGGTCTCCTCTGTCGATATCCGCTTCCGCTAACAGAACTTC
2640 2650 2660 2670 2680 2690 2700
ATGGGACACAGGGCAGTCAACAGTGTGGAAGGGGTCTCCTCTGTCGATATCCGCTTCCGCTAACAGAACTTC
2740 2750 2760 2770 2780 2790 2800
TTGAACACACAGGGGAGAGCCATGCCACCTCTTGAAGAGGTCACTGAAAGAGACATCATTTGCAGAGAAA
2710 2720 2730 2740 2750 2760 2770
TTGAACACACAGGGGAGAGCCATGCCACCTCTTGAAGAGGTCACTGAAAGAGACATCATTTGCAGAGAAA

2810 2820 2830 2840 2850 2860 2870 2880
CTCCTGTCTCAACCCAGAGCTTACAGAGGGTAAAGATCCCATGACGACATGGTCCACGAGTGAAGTGATT
2780 2790 2800 2810 2820 2830 2840
CTCCTGTCTCAACCCAGAGCTTACAGAGGGTAAAGATCCCATGACGACATGGTCCACGAGTGAAGTGATT
2890 2900 2910 2920 2930 2940 2950
TCACCTCAAGAACTGTGACAGGCACAGAGACCTCAGAGGCTTCCGTACTGAAAGAGTTACCGAAGCATCGG
2850 2860 2870 2880 2890 2900 2910 2920
TCACCTCAAGAACTGTGACAGGCACAGAGACCTCAGAGGCTTCCGTACTGAAAGAGTTACCGAAGCATCGG
2930 2940 2950 2960 2970 2980 2990
GGGCCCAAGAGACCAAGAGATGGTTCGCGCAGTTTCCAGCTGACTGATCCCAGACACCAACAGAGAAAG
2960 2970 2980 2990 3000 3010 3020
GGGCCCAAGAGACCAAGAGATGGTTCGCGCAGTTTCCAGCTGACTGATCCCAGACACCAACAGAGAAAG
3030 3040 3050 3060 3070 3080 3090
CCACCCCACTCAGAGAGTAGAGAGTGGTGTCTAGATACAGAAAGAGAGCGCCAGACGAGGCCATCC
3000 3010 3020 3030 3040 3050 3060
CCACCCCACTCAGAGAGTAGAGAGTGGTGTCTAGATACAGAAAGAGAGCGCCAGACGAGGCCATCC
3100 3110 3120 3130 3140 3150 3160
TCCAAAGCCCTTCAGACAAAGGTGAAAGAGGAGTCCCAGGTGCTGCAACCCAGACTGTGCAGAGAACGGGT
3070 3080 3090 3100 3110 3120 3130
TCCAAAGCCCTTCAGACAAAGGTGAAAGAGGAGTCCCAGGTGCTGCAACCCAGACTGTGCAGAGAACGGGT
3170 3180 3190 3200 3210 3220 3230 3240
CAAAAGCACTGAGAGAGGTGAGAGGAGTAGAGGAGGACTCCGAAGTGTGGTTCGAGAGAAAGAGAGACG
3140 3150 3160 3170 3180 3190 3200
CAAAAGCACTGAGAGAGGTGAGAGGAGTAGAGGAGGACTCCGAAGTGTGGTTCGAGAGAAAGAGAGACG
3250 3260 3270 3280 3290 3300 3310
TTATGCGGAAAGGACCCGTGCAAGAGTGGAGCTGAGCATCTTGACAGGGCTCTGAGACTGGACAGGCTA
3210 3220 3230 3240 3250 3260 3270 3280
TTATGCGGAAAGGACCCGTGCAAGAGTGGAGCTGAGCATCTTGACAGGGCTCTGAGACTGGACAGGCTA
3320 3330 3340 3350 3360 3370 3380
CTCCAGAGAGCCTTGAAGTTCCTGAAGTCAACGGCAGATGTAGACCATGTGCCCATGTGCAGGTTATCAAGC
3290 3300 3310 3320 3330 3340 3350
CTCCAGAGAGCCTTGAAGTTCCTGAAGTCAACGGCAGATGTAGACCATGTGCCCATGTGCAGGTTATCAAGC
3390 3400 3410 3420 3430 3440 3450
TCCAGCAGCTGATGGAACAGCGCCGTGGCCCTGAGTCAATCCGAAACCTTTGACAGACAGTGAGACAAATGGAA
3360 3370 3380 3390 3400 3410 3420
TCCAGCAGCTGATGGAACAGCGCCGTGGCCCTGAGTCAATCCGAAACCTTTGACAGACAGTGAGACAAATGGAA
3460 3470 3480 3490 3500 3510 3520
GCACTCCCTTACAGATTACAGACTGAGATGGGACACAGCAAGATGAAACCATTTGACAGCAGGACAGTA
3430 3440 3450 3460 3470 3480 3490
GCACTCCCTTACAGATTACAGACTGAGATGGGACACAGCAAGATGAAACCATTTGACAGCAGGACAGTA
3530 3540 3550 3560 3570 3580 3590 3600
AAGCCACTGACGTGTGAGGAGTCAAGGTCACAGAGAGAGGCGGCTCTGCTCAGAAAGAGAGGCGCTT
3500 3510 3520 3530 3540 3550 3560
AAGCCACTGACGTGTGAGGAGTCAAGGTCACAGAGAGAGGCGGCTCTGCTCAGAAAGAGAGGCGCTT
3610 3620 3630 3640 3650 3660 3670
CGACACTACTTAATATGTTCCAGGCCAGAGAAATCGGGAAGAACCCAGGAAGAGATGTTCTTGAACCTA
3570 3580 3590 3600 3610 3620 3630 3640
CGACACTACTTAATATGTTCCAGGCCAGAGAAATCGGGAAGAACCCAGGAAGAGATGTTCTTGAACCTA
3680 3690 3700 3710 3720 3730 3740
CGACACTACTTAATATGTTCCAGGCCAGAGAAATCGGGAAGAACCCAGGAAGAGATGTTCTTGAACCTA

[illegible]

2. US-09-902-432-3 (1-6160)
 US-08-978-277A-1 Sequence 1, Application US/08978277A

GENERAL INFORMATION:
 APPLICANT: Gelman, Irwin H.
 TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
 NUMBER OF SEQUENCES: 20
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Brunbaugh, Graves, Donohue & Raymond
 STREET: 30 Rockefeller Plaza

4610 4620 4630 4640 4650 4660 4670 4680 4690
 AAGGCAACAAGTTGTGTGTAAGCTTCAATCAGACAGCCGTTGACCGATTCCGACGTTACAGAAACAGCGCCCG
 4510 4520 4530 4540 4550 4560 4570
 AAATCATGCTTATGATTCAACAGACCCAGGTTCTTGATGACAGGCTTGACACAGAGGAGCCCAAGATGCT
 AAATCATGCTTATGATTCAACAGACCCAGGTTCTTGATGACAGGCTTGACACAGAGGAGCCCAAGATGCT
 4480 4490 4460 4470 4480 4490 4500
 GGACAAAATGAAATATCCCAAGATGAAACACCCAGTGCACGACCCCAAGAGAGATTGCAAGTCTTGACCG
 GGACAAAATGAAATATCCCAAGATGAAACACCCAGTGCACGACCCCAAGAGAGATTGCAAGTCTTGACCG
 4350 4360 4370 4380 4390 4400 4410 4420 4430 4440 4450
 TTCTGAGAGCATGAGGCTCAGCTCAGCTGCGAAATGCTTGGCCGCTTGACAGTTGAAAGGCGCGGTGTCAAATGA
 TTCTGAGAGCATGAGGCTCAGCTCAGCTGCGAAATGCTTGGCCGCTTGACAGTTGAAAGGCGCGGTGTCAAATGA
 4220 4230 4240 4250 4260 4270 4280 4290 4300 4310 4320
 CATTTGAGAGCTGCTCTCTCAACCCAAAGATCCAAAGAGAGAGCATGCTGCTGATGAGCCCTCAGCTCAAAGCT
 CATTTGAGAGCTGCTCTCTCAACCCAAAGATCCAAAGAGAGAGCATGCTGCTGATGAGCCCTCAGCTCAAAGCT
 4090 4100 4110 4120 4130 4140 4150 4160 4170 4180 4190
 TAGCCAGGCAAGGCGCATGCTCTCAACCCAAAGATCCAAAGAGAGAGCATGCTGCTGATGAGCCCTCAGCTCAAAGCT
 TAGCCAGGCAAGGCGCATGCTCTCAACCCAAAGATCCAAAGAGAGAGCATGCTGCTGATGAGCCCTCAGCTCAAAGCT
 3860 3870 3880 3890 3900 3910 3920 3930 3940 3950 3960
 CCAGAGAGGCGCATGCTCTCAACCCAAAGATCCAAAGAGAGAGCATGCTGCTGATGAGCCCTCAGCTCAAAGCT
 CCAGAGAGGCGCATGCTCTCAACCCAAAGATCCAAAGAGAGAGCATGCTGCTGATGAGCCCTCAGCTCAAAGCT
 3730 3740 3750 3760 3770 3780 3790 3800 3810 3820 3830
 AGGCCAGGCAAGGCGCATGCTCTCAACCCAAAGATCCAAAGAGAGAGCATGCTGCTGATGAGCCCTCAGCTCAAAGCT
 AGGCCAGGCAAGGCGCATGCTCTCAACCCAAAGATCCAAAGAGAGAGCATGCTGCTGATGAGCCCTCAGCTCAAAGCT
 3600 3610 3620 3630 3640 3650 3660 3670 3680 3690 3700
 GATTCAGAGCCATGAGATCCAGAGAGAGGAGGCGCTCCATATATTTT
 GATTCAGAGCCATGAGATCCAGAGAGAGGAGGCGCTCCATATATTTT
 3470 3480 3490 3500 3510 3520 3530 3540 3550 3560 3570
 TGTACTCTGTAAAGCCGAATGTGAAACAGATCAGAGAAACAAGTCTGCTGTTGGACCTTGAGACCA
 TGTACTCTGTAAAGCCGAATGTGAAACAGATCAGAGAAACAAGTCTGCTGTTGGACCTTGAGACCA
 3340 3350 3360 3370 3380 3390 3400 3410 3420 3430 3440
 GATTTCAGAGCCATGAGATCCAGAGAGAGGAGGCGCTCCATATATTTT
 GATTTCAGAGCCATGAGATCCAGAGAGAGGAGGCGCTCCATATATTTT
 3210 3220 3230 3240 3250 3260 3270 3280 3290 3300 3310
 TGTAGAGCCATGAGATCCAGAGAGAGGAGGCGCTCCATATATTTT
 TGTAGAGCCATGAGATCCAGAGAGAGGAGGCGCTCCATATATTTT
 3080 3090 3100 3110 3120 3130 3140 3150 3160 3170 3180
 GATTTCAGAGCCATGAGATCCAGAGAGAGGAGGCGCTCCATATATTTT
 GATTTCAGAGCCATGAGATCCAGAGAGAGGAGGCGCTCCATATATTTT
 2950 2960 2970 2980 2990 3000 3010 3020 3030 3040 3050
 TGTAGAGCCATGAGATCCAGAGAGAGGAGGCGCTCCATATATTTT
 TGTAGAGCCATGAGATCCAGAGAGAGGAGGCGCTCCATATATTTT
 2820 2830 2840 2850 2860 2870 2880 2890 2900 2910 2920
 GATTTCAGAGCCATGAGATCCAGAGAGAGGAGGCGCTCCATATATTTT
 GATTTCAGAGCCATGAGATCCAGAGAGAGGAGGCGCTCCATATATTTT
 2690 2700 2710 2720 2730 2740 2750 2760 2770 2780 2790
 TGTAGAGCCATGAGATCCAGAGAGAGGAGGCGCTCCATATATTTT
 TGTAGAGCCATGAGATCCAGAGAGAGGAGGCGCTCCATATATTTT
 2560 2570 2580 2590 2600 2610 2620 2630 2640 2650 2660
 TGTAGAGCCATGAGATCCAGAGAGAGGAGGCGCTCCATATATTTT
 TGTAGAGCCATGAGATCCAGAGAGAGGAGGCGCTCCATATATTTT
 2430 2440 2450 2460 2470 2480 2490 2500 2510 2520 2530
 TGTAGAGCCATGAGATCCAGAGAGAGGAGGCGCTCCATATATTTT
 TGTAGAGCCATGAGATCCAGAGAGAGGAGGCGCTCCATATATTTT
 2300 2310 2320 2330 2340 2350 2360 2370 2380 2390 2400
 TGTAGAGCCATGAGATCCAGAGAGAGGAGGCGCTCCATATATTTT
 TGTAGAGCCATGAGATCCAGAGAGAGGAGGCGCTCCATATATTTT
 2170 2180 2190 2200 2210 2220 2230 2240 2250 2260 2270
 TGTAGAGCCATGAGATCCAGAGAGAGGAGGCGCTCCATATATTTT
 TGTAGAGCCATGAGATCCAGAGAGAGGAGGCGCTCCATATATTTT
 2040 2050 2060 2070 2080 2090 2100 2110 2120 2130 2140
 TGTAGAGCCATGAGATCCAGAGAGAGGAGGCGCTCCATATATTTT
 TGTAGAGCCATGAGATCCAGAGAGAGGAGGCGCTCCATATATTTT
 1910 1920 1930 1940 1950 1960 1970 1980 1990 2000 2010
 TGTAGAGCCATGAGATCCAGAGAGAGGAGGCGCTCCATATATTTT
 TGTAGAGCCATGAGATCCAGAGAGAGGAGGCGCTCCATATATTTT
 1780 1790 1800 1810 1820 1830 1840 1850 1860 1870 1880
 TGTAGAGCCATGAGATCCAGAGAGAGGAGGCGCTCCATATATTTT
 TGTAGAGCCATGAGATCCAGAGAGAGGAGGCGCTCCATATATTTT
 1650 1660 1670 1680 1690 1700 1710 1720 1730 1740 1750
 TGTAGAGCCATGAGATCCAGAGAGAGGAGGCGCTCCATATATTTT
 TGTAGAGCCATGAGATCCAGAGAGAGGAGGCGCTCCATATATTTT
 1520 1530 1540 1550 1560 1570 1580 1590 1600 1610 1620
 TGTAGAGCCATGAGATCCAGAGAGAGGAGGCGCTCCATATATTTT
 TGTAGAGCCATGAGATCCAGAGAGAGGAGGCGCTCCATATATTTT
 1390 1400 1410 1420 1430 1440 1450 1460 1470 1480 1490
 TGTAGAGCCATGAGATCCAGAGAGAGGAGGCGCTCCATATATTTT
 TGTAGAGCCATGAGATCCAGAGAGAGGAGGCGCTCCATATATTTT
 1260 1270 1280 1290 1300 1310 1320 1330 1340 1350 1360
 TGTAGAGCCATGAGATCCAGAGAGAGGAGGCGCTCCATATATTTT
 TGTAGAGCCATGAGATCCAGAGAGAGGAGGCGCTCCATATATTTT
 1130 1140 1150 1160 1170 1180 1190 1200 1210 1220 1230
 TGTAGAGCCATGAGATCCAGAGAGAGGAGGCGCTCCATATATTTT
 TGTAGAGCCATGAGATCCAGAGAGAGGAGGCGCTCCATATATTTT
 1000 1010 1020 1030 1040 1050 1060 1070 1080 1090 1100
 TGTAGAGCCATGAGATCCAGAGAGAGGAGGCGCTCCATATATTTT
 TGTAGAGCCATGAGATCCAGAGAGAGGAGGCGCTCCATATATTTT
 870 880 890 900 910 920 930 940 950 960 970
 TGTAGAGCCATGAGATCCAGAGAGAGGAGGCGCTCCATATATTTT
 TGTAGAGCCATGAGATCCAGAGAGAGGAGGCGCTCCATATATTTT
 740 750 760 770 780 790 800 810 820 830 840
 TGTAGAGCCATGAGATCCAGAGAGAGGAGGCGCTCCATATATTTT
 TGTAGAGCCATGAGATCCAGAGAGAGGAGGCGCTCCATATATTTT
 610 620 630 640 650 660 670 680 690 700 710
 TGTAGAGCCATGAGATCCAGAGAGAGGAGGCGCTCCATATATTTT
 TGTAGAGCCATGAGATCCAGAGAGAGGAGGCGCTCCATATATTTT
 480 490 500 510 520 530 540 550 560 570

1170	1180	1190	1200	1210	1220	1230	1240	1250	1260	1270	1280	1290	1300	1310	1320	1330	1340	1350	1360	1370	1380	1390	1400	1410	1420	1430	1440	1450	1460	1470	1480	1490	1500	1510	1520	1530	1540	1550	1560	1570	1580	1590	1600	1610	1620	1630	1640	1650	1660	1670	1680	1690	1700	1710	1720	1730	1740	1750	1760	1770	1780	1790	1800	1810	1820	1830	1840	1850	1860	1870	1880	1890	1900	1910	1920	1930	1940	1950	1960	1970	1980	1990	2000	2010	2020	2030	2040	2050	2060	2070	2080	2090	2100	2110	2120	2130	2140	2150	2160	2170	2180	2190	2200	2210	2220	2230	2240	2250	2260	2270	2280	2290	2300	2310	2320	2330	2340	2350	2360	2370	2380	2390	2400	2410	2420	2430	2440	2450	2460	2470	2480	2490	2500	2510	2520	2530	2540	2550	2560	2570	2580	2590	2600	2610	2620	2630	2640	2650	2660	2670	2680	2690	2700	2710	2720	2730	2740	2750	2760	2770	2780	2790	2800	2810	2820	2830	2840	2850	2860	2870	2880	2890	2900	2910	2920	2930	2940	2950	2960	2970	2980	2990	3000	3010	3020	3030	3040	3050	3060	3070	3080	3090	3100	3110	3120	3130	3140	3150	3160	3170	3180	3190	3200	3210	3220	3230	3240	3250	3260	3270	3280	3290	3300	3310	3320	3330	3340	3350	3360	3370	3380	3390	3400	3410	3420	3430	3440	3450	3460	3470	3480	3490	3500	3510	3520	3530	3540	3550	3560	3570	3580	3590	3600	3610	3620	3630	3640	3650	3660	3670	3680	3690	3700	3710	3720	3730	3740	3750	3760	3770	3780	3790	3800	3810	3820	3830	3840	3850	3860	3870	3880	3890	3900	3910	3920	3930	3940	3950	3960	3970	3980	3990	4000	4010	4020	4030	4040	4050	4060	4070	4080	4090	4100	4110	4120	4130	4140	4150	4160	4170	4180	4190	4200	4210	4220	4230	4240	4250	4260	4270	4280	4290	4300	4310	4320	4330	4340	4350	4360	4370	4380	4390	4400	4410	4420	4430	4440	4450	4460	4470	4480	4490	4500	4510	4520	4530	4540	4550	4560	4570	4580	4590	4600	4610	4620	4630	4640	4650	4660	4670	4680	4690	4700	4710	4720	4730	4740	4750	4760	4770	4780	4790	4800	4810	4820	4830	4840	4850	4860	4870	4880	4890	4900	4910	4920	4930	4940	4950	4960	4970	4980	4990	5000
------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

[illegible]

CCTCAACCCAAAGATCCAAAGGAGCATGCTGCTGATGGCCCTCAGCTCCAAAGCTTAGCCAGGCAGAGGG
 3830 3840 3850 3860 3870 3880 3890
 CAGTGCCTCTGGAAACCTTAACAAAGAATCCCAGACACCAACGACCAAGCTTAACCGAGAGGGCAT-C
 CTTCAACCCCAAAGAT-CAAAAGGAGGATGCTGCTGATGGCCCTCAGCTCCAAAGCTTAGCCAGGCAGAGG-
 CAGTG--TCTGGAAACCTTAACCAAGAATCCCAGACACCAACGACCAAGCTTAACCGAGGA--GCGATGC
 3900 3910 3920 3930 3940 3950
 4990 5000 5010 5020 5030 5040 5050
 CCCCAAGATTGAGTCCAGGAAGAATAATGCTATCAAGTCACTCAAGAGAACCAAGGCCAGGAGAGAG
 CCCCAGAAAGTTGAGTCCAGGAAGAATAATGCTATCAAGTCACTCAAGAGAACCAAGGCCAGGAGAGAG
 3960 3970 3980 3990 4000 4010 4020 4030
 5060 5070 5080 5090 5100 5110 5120
 AGGACCTCGAGGAGCCNAAGGGAGACCTGGAGAATCCTAAGATGTTAGTTGCTCATTTGATCTCTAAGA
 AGGACCTCGAGGAGCCNAAGGGAGACCTGGAGAATCCTAAGATGTTAGTTGCTCATTTGATCTCTAAGA
 4040 4050 4060 4070 4080 4090 4100 4110
 5130 5140 5150 5160 5170 5180 5190
 CCAGATGTGAAACCAAGTCAACGAACAAGATGCTGTTGGGACCTTGAGACCAAGATTTTCAGAGCCCAT
 CCAGATGTGAAACCAAGTCAACGAACAAGATGCTGTTGGGACCTTGAGACCAAGATTTTCAGAGCCCAT
 4110 4120 4130 4140 4150 4160 4170
 5200 5210 5220 5230 5240 5250 5260 5270
 GAGATTCAGAGAGAGGGGCGCTCAATGATTTCCACCAGTAGAGCACCCCGCAAAATTTCTGAGGCTTCATCG
 GAGATTCAGAGAGAGGGGCGCTCAATGATTTCCACCAGTAGAGCACCCCGCAAAATTTCTGAGGCTTCATCG
 4180 4190 4200 4210 4220 4230 4240
 5280 5290 5300 5310 5320 5330 5340
 GGAGCTAGAGCCAGCTAAACATTTCTCGTTTCAAGAGCTGCCCTTGATTGGCCCTTGATGCCGCTCGGTGAT
 GGAGCTAGAGCCAGCTAAACATTTCTCGTTTCAAGAGCTGCCCTTGATTGGCCCTTGATGCCGCTCGGTGAT
 4250 4260 4270 4280 4290 4300 4310
 5350 5360 5370 5380 5390 5400
 TTC--TAAACATTTCT-CGTT-TCAGAGCTGCCCTTGATTT-TGCCCTTGATGCC--GTCCGCTGATTTCG
 TTCGATTTAAGGCTCGGTGTTCTCAA-CTTGGAAACCAATTTCTGCC----ATACCTPAGTTCCACAT-TCACA
 4320 4330 4340 4350 4360 4370 4380
 5410 5420 5430 5440 5450 5460 5470
 GATTTAAG-GTCCCTGGTCTTCAACCTGGAACCAATTCGCGCATACCTAGTTCCACATTTCTCAAACTGGAGCA
 AACTGGAGCATCCTCTTTATGATTTT---ATATGATGTTTTATGATGCTCTCTCCCTCGTACCTATTGTA
 4390 4400 4410 4420 4430 4440 4450
 5480 5490 5500 5510 5520 5530 5540
 TCCTCCTTTATGTA---TTTATATGATGTTTTATGATGCTCTCTCTGATCCTATTGTTATTTTCTTCT
 TATT--TTTTTCTAAGTTAAGCACATGCTTTTTTGTTATTTATGCAATATATACGGGTGTGACCATAGCG
 4460 4470 4480 4490 4500 4510 4520
 5550 5560 5570 5580 5590 5600
 AACGTTT---AAGC-ACATGCTT--TTTGTTATTATGC-----AATATATAACGGGTGT-GCAGCATAG-CGA
 ACGTTTGGAAAGCTCCAAGCCTCAACTGTAAACCTGCAGCAACACAGATACATTTCTTGGCAGAAAGACAA
 4530 4540 4550 4560 4570 4580 4590
 5610 5620 5630 5640 5650 5660
 CGGT-TTGAAGAGCT--CCAAGCCT--CAAAGTGAACCTG-AG-----CAAACAGATAACATCTCTCGGCA
 GTCTTTTAAAGTTTACTGATGCTTAGATCTGTGGGCTTCTAGTCCCTCTGAAAGTGTTGTTTCTCT---A
 4600 4610 4620 4630 4640 4650 4660
 5670 5680 5690 5700 5710 5720 5730
 AGAAGAGACAAGTCTTTTTTAAAGTTTACTGAATGCTTAGATCTGTGGGCTTCTAGTCC--TCTGAAGATGG

1

800 810 820 830 840 850 860

2. US-09-902-432-4 (1-1596)

1480 1490 1500 1510 1520 1530 1540
ADGESFAIEVLAEPEILLESKSNKIYVNIQTAVDQFARTETAPETAYISQTVPCRIUDSREPNC
ADGESFAIEVLAEPEILLESKSNKIYVNIQTAVDQFARTETAPETAYISQTVPCRIUDSREPNC
ADGESFAIEVLAEPEILLESKSNKIYVNIQTAVDQFARTETAPETAYISQTVPCRIUDSREPNC
1100 1110 1120 1130 1140 1150 1160
WTGKAKMKGHPVQPREDLQVLTVLEEM-----LSSEMIALAVESAGVKSIEKLPQPKDKQKHAADG
1170 1180 1190 1200 1210 1220
POLQSLAQAENVSGNLTKESPDTNGPKITEERC
1230 1240 1250 1260

3. US-09-902-432-4 (1-1596)
US-08-978-277A-1 Sequence 18, Application US/08978277A

Sequence 18, Application US/08978277A

GENERAL INFORMATION:
APPLICANT: Gelman, Irwin H.
TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10112-0228
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/978,277A
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/665,401
FILING DATE: 18-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Richard S
REGISTRATION NUMBER: 26,154
REFERENCE/DOCKET NUMBER: A30558 - 165/34008
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-408-2558
TELEFAX: 212-765-2519
TELEX:
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:

Initial Score = 23 Optimized Score = 23 Significance = -0.28
Residue Identity = 100% Matches = 23 Mismatches = 0
Gaps = 0 Conservative Substitutions = 0

700 710 720 730 740 750 760
PRTLGDSHRAEASKDKEAGTDAVPASTQEDQAGSSSPRAGSPSEGEVSTWESFKRLVTPRKKSXK
EGVSTWESFKRLVTPRKKSXK
X 10 20

X 770 780 790 800 810
LEEKAEDSVVQLSTELIEPRRESWVSIIKFIPIGRKKRADGKQEQATVED
L
X

4. US-09-902-432-4 (1-1596)
US-08-978-277A-1 Sequence 17, Application US/08978277A

Sequence 17, Application US/08978277A

GENERAL INFORMATION:
APPLICANT: Gelman, Irwin H.
TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10112-0228
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/978,277A
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/665,401
FILING DATE: 18-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Richard S
REGISTRATION NUMBER: 26,154
REFERENCE/DOCKET NUMBER: A30558 - 165/34008
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-408-2558
TELEFAX: 212-765-2519
TELEX:
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:

Initial Score = 23 Optimized Score = 23 Significance = -0.28
Residue Identity = 100% Matches = 23 Mismatches = 0
Gaps = 0 Conservative Substitutions = 0

550 560 570 580 590 X 600
PESADEQKGESSASSPPEPBTCTCLEKGPLAPQDGEAEGETTSDGSKKGITPMASFKKMTTPKKRVRRP
EGITPMASFKKMTTPKKRVRRP
X 10 20

X 620 630 640 650 660
SESKKEBELKVKSATLSTDSYVSEMDVKTVGEEQKPEEPKRRVDTSV
S
X

5. US-09-902-432-4 (1-1596)
US-08-978-277A-1 Sequence 16, Application US/08978277A

Sequence 16, Application US/08978277A

GENERAL INFORMATION:
APPLICANT: Gelman, Irwin H.
TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: NY
COUNTRY: USA

ZIP: 10112-0228
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/978,277A
FILING DATE:

CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/665,401
FILING DATE: 18-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Richard S
REGISTRATION NUMBER: 26,154
REFERENCE/DOCKET NUMBER: A30558 - 165/34008
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-408-2558
TELEFAX: 212-765-2519
TELEX:

INFORMATION FOR SEQ ID NO: 16:

SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:

Initial Score = 23 Optimized Score = 23 Significance = -0.28
Residue Identity = 100% Matches = 23 Mismatches = 0
Gaps = 0 Conservative Substitutions = 0

460 470 480 490 500 X 510 520
EMCVSGDHTQLTDLSPKPTLPKHPEGIVSEVEMLSQRIKVGSPPLKLFSSGLKSLGKQKGRGG
|||||
KLFSSGLKSLGKQKGRGG
X 10 20

X 530 540 550 560 570
GGDEEPGEYQHITESPESADEQKGSSASSPEPEPTTCLKGPLEAPQD
G
X

6. US-09-902-432-4 (1-1596)
US-08-978-277A-1 Sequence 15, Application US/08978277A

Sequence 15, Application US/08978277A

GENERAL INFORMATION:
APPLICANT: Gelman, Irwin H.
TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: NY
COUNTRY: USA

CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10112-0228
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/978,277A
FILING DATE:

CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/665,401
FILING DATE: 18-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Richard S
REGISTRATION NUMBER: 26,154
REFERENCE/DOCKET NUMBER: A30558 - 165/34008
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-408-2558
TELEFAX: 212-765-2519
TELEX:

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:

Initial Score = 13 Optimized Score = 22 Significance = -0.31
Residue Identity = 76% Matches = 23 Mismatches = 0
Gaps = 7 Conservative Substitutions = 0

230 240 250 260 270 280 290 300
TLKQESSTEIPQAESDQAEEAKDEGEKQKEPTKSPSPSPSSPKKFTTHGWAGWRKKT
|||||
FKSKEDDLETAEKQEAKEKVDDEEKEKTEPASEEQEPAEDTDQARLUSADYEKVEL
20 X

7. US-09-902-432-4 (1-1596)
US-08-978-277A-7 Sequence 7, Application US/08978277A

Sequence 7, Application US/08978277A

GENERAL INFORMATION:
APPLICANT: Gelman, Irwin H.
TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: NY
COUNTRY: USA

ZIP: 10112-0228
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ Version 1.5

Initial Score = 7 Optimized Score = 7 Significance = -0.32
Residue Identity = 100% Matches = 7 Mismatches = 0
Gaps = 0 Conservative Substitutions

250 260 270 280 290 X 300 310
QAESDQAABEAKDEGEKQKPTKSPSPVNSSETTSFKFFTHGWAGWRKKTSTFKSKDEDDLETAE
|||||
WAGWRK
X X

320 330 340
KRKEQAEKVDSEKEKTEPASBEQEPADTDQAR

10. US-09-902-432-4 (1-1596)
US-08-978-277A-1 Sequence 13, Application US/08978277A

Sequence 13, Application US/08978277A

GENERAL INFORMATION:
APPLICANT: Gelman, Irwin H.
TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: NY

COUNTRY: USA
ZIP: 10112-0228
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/978,277A
FILING DATE:

CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/665,401
FILING DATE: 18-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Richard S
REGISTRATION NUMBER: 26,154
REFERENCE/DOCKET NUMBER: A30558 - 165/34008
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-408-2558
TELEFAX: 212-765-2519
TELEX:

INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:

Initial Score = 7 Optimized Score = 8 Significance = -0.32
Residue Identity = 47% Matches = 8 Mismatches = 9
Gaps = 0 Conservative Substitutions

460 470 480 490 500 X 510 520
EMCVSGDHTQLTDLSPBEKTLPKHPGIVSEVMSLSQRIKVGSPPLKLFSSSGLKKLSGKQKXRGG
|||||
KKFSSKKPKLSGFSR
X 10 X

530 540 550 560 570
GGDEEPGEYQHHTSPESADEQKGSASSPPEPTTCKLEKGP

11. US-09-902-432-4 (1-1596)
US-08-978-277A-1 Sequence 12, Application US/08978277A

Sequence 12, Application US/08978277A

GENERAL INFORMATION:
APPLICANT: Gelman, Irwin H.
TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: NY

COUNTRY: USA
ZIP: 10112-0228
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/978,277A
FILING DATE:

CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/665,401
FILING DATE: 18-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Richard S
REGISTRATION NUMBER: 26,154
REFERENCE/DOCKET NUMBER: A30558 - 165/34008
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-408-2558
TELEFAX: 212-765-2519
TELEX:

INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:

Initial Score = 7 Optimized Score = 10 Significance = -0.32
Residue Identity = 47% Matches = 11 Mismatches = 9
Gaps = 3 Conservative Substitutions

460 470 480 490 500 X 510 520
EMCVSGDHTQLTDLSPBEKTLPKHPGIVSEVMSLSQRIKVGSPPLKLFSSSGLKKLSG---KKQKGG
|||||
KKFSSKKPKLSGFSFKSKKE
X 10 20

X 530 540 550 560 570
RGGGDEEPGEYQHHTSPESADEQKGSASSPPEPTTCKLEKGPLEA

A
X

12. US-09-902-432-4 (1-1596)
US-08-978-277A-1 Sequence 11, Application US/08978277A

Sequence 11, Application US/08978277A

GENERAL INFORMATION:
APPLICANT: Gelman, Irwin H.
TITLE OF INVENTION: TUMOR SUPPRESSOR GENE

NAME: Clark, Richard S
REGISTRATION NUMBER: 26,154
REFERENCE/DOCKET NUMBER: A30558 - 165/34008
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-408-2558
TELEFAX: 212-765-2519
TELEX:

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:

Initial Score = 5 Optimized Score = 5 Significance = -0.33
Residue Identity = 50% Matches = 5 Mismatches = 5
Gaps = 0 Conservative Substitutions = 0

250 260 270 280 290 300 320
ABEEAKGEKEKEPTKSPESPSPVNSETTSFKKFTTHGAGWKKTSPKSKEDDLLETAEKRKQEA
|||||
MGCIKSKEDK
X 10

330 340 350
EKVDEBEKEKTEPASEEQEPADTDQARLSADYKVEL

15. US-09-902-432-4 (1-1596)
US-08-978-277A-5 Sequence 5, Application US/08978277A

Sequence 5, Application US/08978277A

GENERAL INFORMATION:
APPLICANT: Gelman, Irwin H.
TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10112-0228
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/978,277A
FILING DATE:

CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/665,401
FILING DATE: 18-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Richard S
REGISTRATION NUMBER: 26,154
REFERENCE/DOCKET NUMBER: A30558 - 165/34008
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-408-2558
TELEFAX: 212-765-2519
TELEX:

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:

Initial Score = 4 Optimized Score = 4 Significance = -0.33
Residue Identity = 40% Matches = 4 Mismatches = 6
Gaps = 0 Conservative Substitutions = 0

710 720 730 740 750 760 770
ASKDKRAGTDAVPASTQEQDQAGSSSPGAGSPGEGGVSTWESFKRLVTRKKSLEKAEADSSVEQL
|||||
MGSSKSKPKD
X 10

780 790 800 810
STEIEPSREESNVSIKKFIPGRKKRADGKQEQATVED

16. US-09-902-432-4 (1-1596)
US-08-978-277A-1 Sequence 14, Application US/08978277A

Sequence 14, Application US/08978277A

GENERAL INFORMATION:
APPLICANT: Gelman, Irwin H.
TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10112-0228
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/978,277A
FILING DATE:

CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/665,401
FILING DATE: 18-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Richard S
REGISTRATION NUMBER: 26,154
REFERENCE/DOCKET NUMBER: A30558 - 165/34008
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-408-2558
TELEFAX: 212-765-2519
TELEX:

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:

Initial Score = 4 Optimized Score = 4 Significance = -0.33
Residue Identity = 20% Matches = 4 Mismatches = 16
Gaps = 0 Conservative Substitutions = 0

690 700 710 720 730 740 750 X

SSSDDEGPTLTGDSHRAEASDKKAGTDVAVPASTQEDQAQSSSPSESGVSTWESFKRLVT

KRRMKAFIAVSAARFKC
X 10 20

760 770 780 790 800
PRKSKSLKEKAEDSSVEQLSTIEFRESWSVSIKKFIPGRKKRA

17. US-09-902-432-4 (1-1596)
US-08-978-277A-9 Sequence 9, Application US/08978277A

Sequence 9, Application US/08978277A
GENERAL INFORMATION:

APPLICANT: Gelman, Irwin H.
TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10112-0228

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/978,277A

CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/665,401
FILING DATE: 18-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Richard S
REGISTRATION NUMBER: 26,154
REFERENCE/DOCKET NUMBER: A30558 - 165/34008
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-408-2558
TELEFAX: 212-765-2519

TELEX:
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:

Initial Score = 3 Optimized Score = 4 Significance = -0.33
Residue Identity = 40% Matches = 4 Mismatches = 6
Gaps = 0 Conservative Substitutions = 0

580 590 600 610 620 630 640
DGEAEGTTSDEGKEREGITTPWASFKKAVTPRPSBSDEKELEVKASATLSSTDSTVSEMODEVKTV
|||
MGCTLSAEDK
X 10

650 660 670 680
GGEOKPEEPKRRVDTSVSWALICVGSSEKRRARAKASS

18. US-09-902-432-4 (1-1596)
US-08-978-277A-1 Sequence 10, Application US/08978277A

Sequence 10, Application US/08978277A

GENERAL INFORMATION:

APPLICANT: Gelman, Irwin H.
TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10112-0228

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/978,277A
FILING DATE:

CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/665,401
FILING DATE: 18-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Richard S
REGISTRATION NUMBER: 26,154
REFERENCE/DOCKET NUMBER: A30558 - 165/34008
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-408-2558
TELEFAX: 212-765-2519

TELEX:
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:

Initial Score = 2 Optimized Score = 2 Significance = -0.34
Residue Identity = 20% Matches = 2 Mismatches = 8
Gaps = 0 Conservative Substitutions = 0

560 570 580 590 600 610 620
SASSPEPEETTCLEKGPLAPDGEAEBEGTTSDEKKRGITTPWASFKKAVTPRPSBSDEKELEK
|||
MLCCMRITQ
X 10

630 640 650 660
VKSAATLSSTDSTVSEMODEVKTVGEOKPEEPKRRVDT

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 12, 2002, 14:13:30 ; Search time 11503 Seconds
(without alignments)
13464.116 Million cell updates/sec

Title: US-09-902-432-3
Perfect score: 6160
Sequence: 1 cgggcccggggaggaag.....aataaacattattgactac 6160

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 24791104 seqs, 12571243825 residues

Total number of hits satisfying chosen parameters: 49582208

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents NA Main.*
1: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq.*
2: /cgn2_6/ptodata/2/pna/US06_COMB.seq.*
3: /cgn2_6/ptodata/2/pna/US07_COMB.seq.*
4: /cgn2_6/ptodata/2/pna/US08_COMB.seq.*
5: /cgn2_6/ptodata/2/pna/US081_COMB.seq.*
6: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
7: /cgn2_6/ptodata/2/pna/US083_COMB.seq.*
8: /cgn2_6/ptodata/2/pna/US084_COMB.seq.*
9: /cgn2_6/ptodata/2/pna/US085_COMB.seq.*
10: /cgn2_6/ptodata/2/pna/US086_COMB.seq.*
11: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
12: /cgn2_6/ptodata/2/pna/US088_COMB.seq.*
13: /cgn2_6/ptodata/2/pna/US089_COMB.seq.*
14: /cgn2_6/ptodata/2/pna/US090_COMB.seq.*
15: /cgn2_6/ptodata/2/pna/US091_COMB.seq.*
16: /cgn2_6/ptodata/2/pna/US092_COMB.seq.*
17: /cgn2_6/ptodata/2/pna/US093_COMB.seq.*
18: /cgn2_6/ptodata/2/pna/US094_COMB.seq.*
19: /cgn2_6/ptodata/2/pna/US095A_COMB.seq.*
20: /cgn2_6/ptodata/2/pna/US095B_COMB.seq.*
21: /cgn2_6/ptodata/2/pna/US095C_COMB.seq.*
22: /cgn2_6/ptodata/2/pna/US095D_COMB.seq.*
23: /cgn2_6/ptodata/2/pna/US096A_COMB.seq.*
24: /cgn2_6/ptodata/2/pna/US096B_COMB.seq.*
25: /cgn2_6/ptodata/2/pna/US096C_COMB.seq.*
26: /cgn2_6/ptodata/2/pna/US096D_COMB.seq.*
27: /cgn2_6/ptodata/2/pna/US096E_COMB.seq.*
28: /cgn2_6/ptodata/2/pna/US097A_COMB.seq.*
29: /cgn2_6/ptodata/2/pna/US097B_COMB.seq.*
30: /cgn2_6/ptodata/2/pna/US097C_COMB.seq.*
31: /cgn2_6/ptodata/2/pna/US098A_COMB.seq.*
32: /cgn2_6/ptodata/2/pna/US098B_COMB.seq.*
33: /cgn2_6/ptodata/2/pna/US098C_COMB.seq.*
34: /cgn2_6/ptodata/2/pna/US099A_COMB.seq.*
35: /cgn2_6/ptodata/2/pna/US099B_COMB.seq.*
36: /cgn2_6/ptodata/2/pna/US099C_COMB.seq.*
37: /cgn2_6/ptodata/2/pna/US099D_COMB.seq.*
38: /cgn2_6/ptodata/2/pna/US100A_COMB.seq.*
39: /cgn2_6/ptodata/2/pna/US100B_COMB.seq.*
40: /cgn2_6/ptodata/2/pna/US101A_COMB.seq.*
41: /cgn2_6/ptodata/2/pna/US101B_COMB.seq.*
42: /cgn2_6/ptodata/2/pna/US102A_COMB.seq.*
43: /cgn2_6/ptodata/2/pna/US102B_COMB.seq.*

44: /cgn2_6/ptodata/2/pna/US6000_COMB.seq.*
45: /cgn2_6/ptodata/2/pna/US6001_COMB.seq.*
46: /cgn2_6/ptodata/2/pna/US6002_COMB.seq.*
47: /cgn2_6/ptodata/2/pna/US6003_COMB.seq.*
48: /cgn2_6/ptodata/2/pna/US6004_COMB.seq.*
49: /cgn2_6/ptodata/2/pna/US6005_COMB.seq.*
50: /cgn2_6/ptodata/2/pna/US6006_COMB.seq.*
51: /cgn2_6/ptodata/2/pna/US6007_COMB.seq.*
52: /cgn2_6/ptodata/2/pna/US6008_COMB.seq.*
53: /cgn2_6/ptodata/2/pna/US6009_COMB.seq.*
54: /cgn2_6/ptodata/2/pna/US6010_COMB.seq.*
55: /cgn2_6/ptodata/2/pna/US6011_COMB.seq.*
56: /cgn2_6/ptodata/2/pna/US6012_COMB.seq.*
57: /cgn2_6/ptodata/2/pna/US6013_COMB.seq.*
58: /cgn2_6/ptodata/2/pna/US6014_COMB.seq.*
59: /cgn2_6/ptodata/2/pna/US6015_COMB.seq.*
60: /cgn2_6/ptodata/2/pna/US6016_COMB.seq.*
61: /cgn2_6/ptodata/2/pna/US6017_COMB.seq.*
62: /cgn2_6/ptodata/2/pna/US6018_COMB.seq.*
63: /cgn2_6/ptodata/2/pna/US6019_COMB.seq.*
64: /cgn2_6/ptodata/2/pna/US6020_COMB.seq.*
65: /cgn2_6/ptodata/2/pna/US6021_COMB.seq.*
66: /cgn2_6/ptodata/2/pna/US6022_COMB.seq.*
67: /cgn2_6/ptodata/2/pna/US6023_COMB.seq.*
68: /cgn2_6/ptodata/2/pna/US6024_COMB.seq.*
69: /cgn2_6/ptodata/2/pna/US6025_COMB.seq.*
70: /cgn2_6/ptodata/2/pna/US6026_COMB.seq.*
71: /cgn2_6/ptodata/2/pna/US6027_COMB.seq.*
72: /cgn2_6/ptodata/2/pna/US6028_COMB.seq.*
73: /cgn2_6/ptodata/2/pna/US6029_COMB.seq.*
74: /cgn2_6/ptodata/2/pna/US6030_COMB.seq.*
75: /cgn2_6/ptodata/2/pna/US6031_COMB.seq.*
76: /cgn2_6/ptodata/2/pna/US6032_COMB.seq.*
77: /cgn2_6/ptodata/2/pna/US6033_COMB.seq.*
78: /cgn2_6/ptodata/2/pna/US6034_COMB.seq.*
79: /cgn2_6/ptodata/2/pna/US6035_COMB.seq.*
80: /cgn2_6/ptodata/2/pna/US6036_COMB.seq.*
81: /cgn2_6/ptodata/2/pna/US6037_COMB.seq.*
82: /cgn2_6/ptodata/2/pna/US6038_COMB.seq.*
83: /cgn2_6/ptodata/2/pna/US6039_COMB.seq.*
84: /cgn2_6/ptodata/2/pna/US6040_COMB.seq.*
85: /cgn2_6/ptodata/2/pna/US6041_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match	Length	DB	ID	Description
	1	6160	100.0	6160	1	PCT-US97-06830-3	Sequence 3, Appli
	2	6160	100.0	6160	10	US-08-665-401-3	Sequence 3, Appli
	3	6160	100.0	6160	34	US-09-902-432-3	Sequence 3, Appli
	4	5204.2	84.5	5236	1	PCT-US02-16173A-2063	Sequence 2063, Ap
	5	5204.2	84.5	5236	41	US-10-191-803-279	Sequence 279, App
	6	5200	84.4	5200	13	US-08-978-277A-3	Sequence 3, Appli
	7	4831.4	78.4	5134	1	PCT-US97-06830-1	Sequence 1, Appli
	8	4831.4	78.4	5134	10	US-08-665-401-1	Sequence 1, Appli
	9	4831.4	78.4	5134	34	US-09-902-432-1	Sequence 1, Appli
	10	4701.4	76.3	5074	13	US-08-978-277A-1	Sequence 1, Appli
	11	4548.6	73.8	6199	80	US-60-360-207-12210	Sequence 12210, A
	12	2181.2	35.4	7651	30	US-09-770-174-4018	Sequence 4018, Ap
	13	2178	35.4	6603	71	US-60-278-232-10143	Sequence 10143, A
c	14	2178	35.4	6607	68	US-60-245-228-517	Sequence 517, App
	15	2172.8	35.3	6632	76	US-60-324-185-30580	Sequence 30580, A
	16	2160.4	35.1	8660	25	US-09-644-867-6364	Sequence 6364, Ap
	17	2160.4	35.1	8660	25	US-09-644-869-9117	Sequence 9117, Ap
	18	2160.4	35.1	8660	25	US-09-652-911-10117	Sequence 10117, A
	19	2160.4	35.1	8660	25	US-09-652-913-10715	Sequence 10715, A
	20	2160.4	35.1	8660	25	US-09-652-918-8758	Sequence 8758, Ap
	21	2160.4	35.1	8660	26	US-09-667-760-2000	Sequence 2000, Ap

22 2160.4 35.1 8660 27 US-09-699-997-11667
 23 2160.4 35.1 8660 27 US-09-699-999-6654
 24 2160.4 35.1 8660 28 US-09-700-000-5679
 25 2160.4 35.1 8660 28 US-09-710-280-2315
 26 2160.4 35.1 8660 28 US-09-710-286-3428
 27 2160.4 35.1 8660 28 US-09-716-472-4599
 28 2160.4 35.1 8660 28 US-09-716-953-2245
 29 2160.4 35.1 8660 29 US-09-726-789-3701
 30 2160.4 35.1 8660 29 US-09-726-802-2407
 31 2160.4 35.1 8660 29 US-09-726-805-1797
 32 2160.4 35.1 8660 29 US-09-726-806-5392
 33 2159.6 35.1 8492 23 US-09-606-776-5120
 34 2144.4 34.8 6608 16 US-09-320-132-58
 35 2144.4 34.8 6608 29 US-09-738-877-1
 36 2144.4 34.8 6608 33 US-09-880-107-3439
 37 2144.4 34.8 6608 41 US-10-155-653-58
 38 2108.8 34.2 8886 76 US-60-324-185-30584
 39 2106.2 34.2 8869 71 US-60-278-232-10145
 40 2106.2 34.2 8888 65 US-60-213-359-7019
 41 2094.6 34.0 8877 61 US-60-172-360-24882
 42 2055.6 33.4 8243 17 US-09-397-022-4579
 43 2055.6 33.4 8243 18 US-09-404-549-4757
 44 2055.6 33.4 8243 18 US-09-404-549A-4757
 45 2055.6 33.4 8243 18 US-09-428-151A-10174

Sequence 11667, A
 Sequence 6654, Ap
 Sequence 5679, Ap
 Sequence 2315, Ap
 Sequence 3428, Ap
 Sequence 4599, Ap
 Sequence 2245, Ap
 Sequence 3701, Ap
 Sequence 2407, Ap
 Sequence 1797, Ap
 Sequence 5392, Ap
 Sequence 5120, Ap
 Sequence 58, Ap
 Sequence 1, Appl
 Sequence 3439, Ap
 Sequence 58, Appl
 Sequence 10145, A
 Sequence 7019, Ap
 Sequence 2488, A
 Sequence 4579, Ap
 Sequence 4757, Ap
 Sequence 10174, A

ALIGNMENTS

RESULT 1
 PCT-US97-06830-3
 Sequence 3, Application PC/TUS9706830
 GENERAL INFORMATION:
 APPLICANT: Gelman, Irwin H.
 TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
 NUMBER OF SEQUENCES: 20
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
 STREET: 30 Rockefeller Plaza
 CITY: New York
 STATE: NY
 COUNTRY: USA
 ZIP: 10112-0228
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq Version 1.5
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US97/06830
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/665,401
 FILING DATE: 18-JUN-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Clark, Richard S
 REGISTRATION NUMBER: 26,154
 REFERENCE/DOCKET NUMBER: A30558 - 165/34008
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-408-2558
 TELEFAX: 212-765-2519
 TELEX:
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 6160 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 FRAGMENT TYPE:

ORIGINAL SOURCE:

PCT-US97-06830-3

Query Match 100.0%; Score 6160; DB 1; Length 6160;
 Best local Similarity 100.0%; Pred. No. 0;
 Matches 6160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGGCCGGCGGGGATGAAAGACCTGAGCCATGGCGCGAGCACTTCCACCGAGAGC 60
 1 CGGGCCGGCGGGGATGAAAGACCTGAGCCATGGCGCGAGCACTTCCACCGAGAGC 60
 QY 61 GGAAGCCCGAGAGCCCGCGGGAGCAGACCGCGAGAGAGTGGTCTGATGGCCATG 120
 61 GGAAGCCCGAGAGCCCGCGGGAGCAGACCGCGAGAGAGTGGTCTGATGGCCATG 120
 Db 61 GGAAGCCCGAGAGCCCGCGGGAGCAGACCGCGAGAGAGTGGTCTGATGGCCATG 120
 QY 121 GGCCCGAGCTGAAAGCCTCGGGAGACCTGAGACCCCGCGAGCCGCGACCCA 180
 121 GGCCCGAGCTGAAAGCCTCGGGAGACCTGAGACCCCGCGAGCCGCGACCCA 180
 Db 121 GGCCCGAGCTGAAAGCCTCGGGAGACCTGAGACCCCGCGAGCCGCGACCCA 180
 QY 181 AGCTCCACAGAGAAATGGCCAGCTGCTTCTGTCAACGCGCTAGCTGAACAGAGATG 240
 181 AGCTCCACAGAGAAATGGCCAGCTGCTTCTGTCAACGCGCTAGCTGAACAGAGATG 240
 Db 181 AGCTCCACAGAGAAATGGCCAGCTGCTTCTGTCAACGCGCTAGCTGAACAGAGATG 240
 QY 241 TCCATGTCGAAGAGAAACCGAGAGGGGAGAGAGAAAGTGTGATGAGGATGTTG 300
 241 TCCATGTCGAAGAGAAACCGAGAGGGGAGAGAGAAAGTGTGATGAGGATGTTG 300
 Db 241 TCCATGTCGAAGAGAAACCGAGAGGGGAGAGAGAAAGTGTGATGAGGATGTTG 300
 QY 301 GACAGCGAGAGTCAGAAATGTGAGAGAAAGACCGAGTTGAAAGATGGCGCACT 360
 301 GACAGCGAGAGTCAGAAATGTGAGAGAAAGACCGAGTTGAAAGATGGCGCACT 360
 Db 301 GACAGCGAGAGTCAGAAATGTGAGAGAAAGACCGAGTTGAAAGATGGCGCACT 360
 QY 361 CCACAGCTGTTGAAGATATCAAAAGATGGCGAGAGAGACATCAAGAAATTAATGAC 420
 361 CCACAGCTGTTGAAGATATCAAAAGATGGCGAGAGAGACATCAAGAAATTAATGAC 420
 Db 361 CCACAGCTGTTGAAGATATCAAAAGATGGCGAGAGAGACATCAAGAAATTAATGAC 420
 QY 421 AGATCCCTGCTTCAGAAACATGTGAAAGAAATGTACAGCCTGTAGTCCAGGCTA 480
 421 AGATCCCTGCTTCAGAAACATGTGAAAGAAATGTACAGCCTGTAGTCCAGGCTA 480
 Db 421 AGATCCCTGCTTCAGAAACATGTGAAAGAAATGTACAGCCTGTAGTCCAGGCTA 480
 QY 481 ATGATGTTGGCTTCAGAAAGATATTAATTTGTTGTTTAAATTCACGGTGAAGAG 540
 481 ATGATGTTGGCTTCAGAAAGATATTAATTTGTTGTTTAAATTCACGGTGAAGAG 540
 Db 481 ATGATGTTGGCTTCAGAAAGATATTAATTTGTTGTTTAAATTCACGGTGAAGAG 540
 QY 541 ATAAATATGAAATATCATATCTGTCCAATCACTCATCTGTCACAAAGATGAAGCGCAAG 600
 541 ATAAATATGAAATATCATATCTGTCCAATCACTCATCTGTCACAAAGATGAAGCGCAAG 600
 Db 541 ATAAATATGAAATATCATATCTGTCCAATCACTCATCTGTCACAAAGATGAAGCGCAAG 600
 QY 601 GGGCAGAGAGCTGTGTGGAGCTGAGACCCAGAGAGCCAGTGTGAGACTGCCGTG 660
 601 GGGCAGAGAGCTGTGTGGAGCTGAGACCCAGAGAGCCAGTGTGAGACTGCCGTG 660
 Db 601 GGGCAGAGAGCTGTGTGGAGCTGAGACCCAGAGAGCCAGTGTGAGACTGCCGTG 660
 QY 661 GAGAGTCAGATCCAAAGAAAGTGAAGCTGAAGCAATCCAGAGAAACCAAGAGCCACC 720
 661 GAGAGTCAGATCCAAAGAAAGTGAAGCTGAAGCAATCCAGAGAAACCAAGAGCCACC 720
 Db 661 GAGAGTCAGATCCAAAGAAAGTGAAGCTGAAGCAATCCAGAGAAACCAAGAGCCACC 720
 QY 721 TGAAGCAAGAAACAGAGCAGCAAGAAATCCCTTCAAGCCGAATCTGATCAAGCGCTG 780
 721 TGAAGCAAGAAACAGAGCAGCAAGAAATCCCTTCAAGCCGAATCTGATCAAGCGCTG 780
 Db 721 TGAAGCAAGAAACAGAGCAGCAAGAAATCCCTTCAAGCCGAATCTGATCAAGCGCTG 780
 QY 781 AGAAGAAAGCCAAAGATGAAGAGGAGAAAGAAACCAAGAAAGAGCCACCAAGTCCCGAG 840
 781 AGAAGAAAGCCAAAGATGAAGAGGAGAAAGAAACCAAGAAAGAGCCACCAAGTCCCGAG 840
 Db 781 AGAAGAAAGCCAAAGATGAAGAGGAGAAAGAAACCAAGAAAGAGCCACCAAGTCCCGAG 840
 QY 841 AATCCCGAGAGCCAGTCAACAGTGAACAATCTTCTTCAAGAAAGTTCTTCACTC 900
 841 AATCCCGAGAGCCAGTCAACAGTGAACAATCTTCTTCAAGAAAGTTCTTCACTC 900
 Db 841 AATCCCGAGAGCCAGTCAACAGTGAACAATCTTCTTCAAGAAAGTTCTTCACTC 900
 QY 901 ACGTTGGGCGGCTGGGCGAAGAACCAAGCTTCAAGAAATCAAAAGAGATATCTGG 960
 901 ACGTTGGGCGGCTGGGCGAAGAACCAAGCTTCAAGAAATCAAAAGAGATATCTGG 960
 Db 901 ACGTTGGGCGGCTGGGCGAAGAACCAAGCTTCAAGAAATCAAAAGAGATATCTGG 960
 QY 961 AAACCTGCCAAGAGAAAGAGAGCAAGAGCAAGAAAGTGAAGAGAAAGAAAGAAA 1020
 961 AAACCTGCCAAGAGAAAGAGAGCAAGAGCAAGAAAGTGAAGAGAAAGAAAGAAA 1020
 Db 961 AAACCTGCCAAGAGAAAGAGAGCAAGAGCAAGAAAGTGAAGAGAAAGAAAGAAA 1020

Db 961 AAACCTGCCGAGAGAGAAAGGAGCAGAGGCGAGAAAGTAGACGAGGAAGAAAGGAA 1020
Qy 1021 AGACAGAGCCAGCCTCGAGAGAGCAGGCGCGGAGAGACACACAGACAGAGCCAGGTTGT 1080
Db 1021 AGACAGAGCCAGCCTCGAGAGAGCAGGAGCGCGGAGAGACACACAGACAGAGCCAGGTTGT 1080
Qy 1081 CAGCAGACTACGAGAGGTGGAGCTGCTCTTTGGCAACGGAAGTGTGATGAGAAGATGGAAGCCC 1140
Db 1081 CAGCAGACTACGAGAGGTGGAGCTGCTCTTTGGCAACGGAAGTGTGATGAGAAGATGGAAGCCC 1140
Qy 1141 CGTCAGAGGAGAGTGTCTCTTTGGCAACGGAAGTGTGATGAGAAGATGGAAGCCC 1200
Db 1141 CGTCAGAGGAGAGTGTCTCTTTGGCAACGGAAGTGTGATGAGAAGATGGAAGCCC 1200
Qy 1201 ACCAAGAGTGTGTCAGAGGTCCACGTGAGCACCGTGAGGAAGACAGAGAGGAGCAGG 1260
Db 1201 ACCAAGAGTGTGTCAGAGGTCCACGTGAGCACCGTGAGGAAGACAGAGAGGAGCAGG 1260
Qy 1261 GAGGAGGAGGAGGCTGAAGGGGCGTGTGTGAGGAAGACAGAGGAGGAGGAGG 1320
Db 1261 GAGGAGGAGGAGGCTGAAGGGGCGTGTGTGAGGAAGACAGAGGAGGAGGAGG 1320
Qy 1321 CTGAGAACTCGCTGAGCCCGCAGAGGTCTCCCGAGGAGCTGAGCCTGCTGAGGAGCTGA 1380
Db 1321 CTGAGAACTCGCTGAGCCCGCAGAGGTCTCCCGAGGAGCTGAGCCTGCTGAGGAGCTGA 1380
Qy 1381 TGAAGAGCAGAGAGATGTGTCTCTGAGGAGGAGCCACACTGACAGACCTAAGTC 1440
Db 1381 TGAAGAGCAGAGAGATGTGTCTCTGAGGAGGAGCCACACTGACAGACCTAAGTC 1440
Qy 1441 CTGAAGAGAGACGCTGCCAAACACCCAGAGAGGATGTGAGTGAAGGAACAGAGATGCTGT 1500
Db 1441 CTGAAGAGAGACGCTGCCAAACACCCAGAGAGGATGTGAGTGAAGGAACAGAGATGCTGT 1500
Qy 1501 CCTCTCAGGAAGAACTCAAGGTACAGGGAAGTCCCTTGAAGAACTCTTCAGTAGCTCAG 1560
Db 1501 CCTCTCAGGAAGAACTCAAGGTACAGGGAAGTCCCTTGAAGAACTCTTCAGTAGCTCAG 1560
Qy 1561 GCTTAAAGAACTGTCTGGGAAGAGCAGAGAGGGAACAGAGGAGGTGGGGAGACGAAG 1620
Db 1561 GCTTAAAGAACTGTCTGGGAAGAGCAGAGAGGGAACAGAGGAGGTGGGGAGACGAAG 1620
Qy 1621 AGCCTGGAGAACTACCAACATTCACCCGAATCCCAAGAGAGTGTGATGAGCAGAGG 1680
Db 1621 AGCCTGGAGAACTACCAACATTCACCCGAATCCCAAGAGAGTGTGATGAGCAGAGG 1680
Qy 1681 GAGAGAGCTCTGCTGCTGCCCGAGGAGCCTGAGGAGACCACTGTCTGAGAAAGGGC 1740
Db 1681 GAGAGAGCTCTGCTGCTGCCCGAGGAGCCTGAGGAGACCACTGTCTGAGAAAGGGC 1740
Qy 1741 CGCTGGAAGCACCCAGGATGGGAAGCTGAGGAAGGAACCTACTTCCGATGGAGAGAAGA 1800
Db 1741 CGCTGGAAGCACCCAGGATGGGAAGCTGAGGAAGGAACCTACTTCCGATGGAGAGAAGA 1800
Qy 1801 AGAGAGAGGAGATCACTCCCTGGGATCCTTCAAAAAGATGGTGACACCCCAAGAAACGGG 1860
Db 1801 AGAGAGAGGAGATCACTCCCTGGGATCCTTCAAAAAGATGGTGACACCCCAAGAAACGGG 1860
Qy 1861 TCCCAAGACCTTCTGAGAGTCACAGGAGGAGAGCTGAGGAAGGTCAAGAGCCCACT 1920
Db 1861 TCCCAAGACCTTCTGAGAGTCACAGGAGGAGAGCTGAGGAAGGTCAAGAGCCCACT 1920
Qy 1921 TGTCTCCACTGATAGCACAGTGTGAGAAATGCAAGATGAAAGTCAAAACTGTGTGAGG 1980
Db 1921 TGTCTCCACTGATAGCACAGTGTGAGAAATGCAAGATGAAAGTCAAAACTGTGTGAGG 1980
Qy 1981 AACAAAAGCCAGAGAAACAAAGCGTAGGTGGATACCTTCAAGTGTCTTGGGAAGCACTGA 2040
Db 1981 AACAAAAGCCAGAGAAACAAAGCGTAGGTGGATACCTTCAAGTGTCTTGGGAAGCACTGA 2040
Qy 2041 TTTGTGTCGGATCATCCAGAGAGAGCAAGAGGATCTCTTTCAGATGATGAAGGAG 2100
Db 2041 TTTGTGTCGGATCATCCAGAGAGAGCAAGAGGATCTCTTTCAGATGATGAAGGAG 2100

Qy 2101 GGCCAAAGGACACTGGGAGGGGACAGTCAACAGACAGAGGAGGCCAGCAAAAGCAAAAG 2160
Db 2101 GGCCAAAGGACACTGGGAGGGGACAGTCAACAGACAGAGGAGGCCAGCAAAAGCAAAAG 2160
Qy 2161 CCGGAAACAGAGCGTGTCTGCCAGCACCCAGGAGCAGGACCAAGCGCAAGGAGTTTCT 2220
Db 2161 CCGGAAACAGAGCGTGTCTGCCAGCACCCAGGAGCAGGACCAAGCGCAAGGAGTTTCT 2220
Qy 2221 CACCCGAGCAGCGGGAAGCCCTTCCGAAGGGGAAGTGTCTCCACTTGGGAGTCAATTA 2280
Db 2221 CACCCGAGCAGCGGGAAGCCCTTCCGAAGGGGAAGTGTCTCCACTTGGGAGTCAATTA 2280
Qy 2281 AAAGATTAGTCACTCCAAAGAAAAAATCCAAGTCAAAATCTGGAGAGAAAAAGCAAGACT 2340
Db 2281 AAAGATTAGTCACTCCAAAGAAAAAATCCAAGTCAAAATCTGGAGAGAAAAAGCAAGACT 2340
Qy 2341 CTAGTGTAGAGCAGTGTCTCACTGAGATCGAAACGAGTAGAGAAATCTTTGGTTTCCA 2400
Db 2341 CTAGTGTAGAGCAGTGTCTCACTGAGATCGAAACGAGTAGAGAAATCTTTGGTTTCCA 2400
Qy 2401 TTAAGAAATTCATCCCGGACGCGGAAGAAAAAGGGCAGACGGAAGCAAGAAACAAAGCCA 2460
Db 2401 TTAAGAAATTCATCCCGGACGCGGAAGAAAAAGGGCAGACGGAAGCAAGAAACAAAGCCA 2460
Qy 2461 CTGTGGAAGACTCAGGGCCAGTGGAGATAAATGAGGACGACCCCTTAATGTCTCCAGCCGTG 2520
Db 2461 CTGTGGAAGACTCAGGGCCAGTGGAGATAAATGAGGACGACCCCTTAATGTCTCCAGCCGTG 2520
Qy 2521 TGCTCTGTCTGAGTATATGTCAGTGGAGAGGAGAGATGGAAGCCAGGGGAAATACGG 2580
Db 2521 TGCTCTGTCTGAGTATATGTCAGTGGAGAGGAGAGATGGAAGCCAGGGGAAATACGG 2580
Qy 2581 AGCTGCCCCAGCTGTCTGGGGCTGTGTACGTGTCCGAGGAGCTCAGTAAGACTCTGTGTC 2640
Db 2581 AGCTGCCCCAGCTGTCTGGGGCTGTGTACGTGTCCGAGGAGCTCAGTAAGACTCTGTGTC 2640
Qy 2641 ACATGTGAGTGTGCGCAGTCAATTGATGGGACACAGGGCAGTCACCCAGTGTGGAAGAGCGGT 2700
Db 2641 ACATGTGAGTGTGCGCAGTCAATTGATGGGACACAGGGCAGTCACCCAGTGTGGAAGAGCGGT 2700
Qy 2701 CTCTCTGCTGGATATCCGCTTCGTTACAGAACTCTTTGAACACACAGCGGGAGAAAGCCA 2760
Db 2701 CTCTCTGCTGGATATCCGCTTCGTTACAGAACTCTTTGAACACACAGCGGGAGAAAGCCA 2760
Qy 2761 TGCCACCTGTGAAGAGGTCACTGAAAAAGACATCATTTGAGAGAACTCTCTGTGCTCA 2820
Db 2761 TGCCACCTGTGTGAAGAGGTCACTGAAAAAGACATCATTTGAGAGAACTCTCTGTGCTCA 2820
Qy 2821 CCCAGACGTTTACCAGAGGGTAAAGATGCCCATGACGACATGGTCCACAGTGAAGTGGATT 2880
Db 2821 CCCAGACGTTTACCAGAGGGTAAAGATGCCCATGACGACATGGTCCACAGTGAAGTGGATT 2880
Qy 2881 TCACCTCAGAACTGTGACAGCCACAGAGACCTCAGAGGCTCTCCGCTACTGAAAGAGTTA 2940
Db 2881 TCACCTCAGAACTGTGACAGCCACAGAGACCTCAGAGGCTCTCCGCTACTGAAAGAGTTA 2940
Qy 2941 CCGAAGCATCGGGGCGGAGAGACCAACATGTTGTCGCGAGTTTCCAGCTGACTG 3000
Db 2941 CCGAAGCATCGGGGCGGAGAGACCAACATGTTGTCGCGAGTTTCCAGCTGACTG 3000
Qy 3001 ACTCCCCAGACACACAGAGGAAGCCACCCAGTTCAGGAGGTAGAGAGTGGTGTGCTAG 3060
Db 3001 ACTCCCCAGACACACAGAGGAAGCCACCCAGTTCAGGAGGTAGAGAGTGGTGTGCTAG 3060
Qy 3061 ATACAGAAAGAGAGGAGCGCCAGACGAGGCCATCTCTCAAGCGCTTGCAGACAAAGTGA 3120
Db 3061 ATACAGAAAGAGAGGAGCGCCAGACGAGGCCATCTCTCAAGCGCTTGCAGACAAAGTGA 3120
Qy 3121 AAGAGGAGTCCCAGGTGCTGCAACCCAGACTGTGCAGAGAACGGGGTCAAAAGCACTGG 3180
Db 3121 AAGAGGAGTCCCAGGTGCTGCAACCCAGACTGTGCAGAGAACGGGGTCAAAAGCACTGG 3180

QY	3181	AGAAAGTTGAGAGAGTGAAGAGAGACTCCGAAAGTGTGGCTTCCGAGAAAGAGAGACG	3240
Db	3181	AGAAAGTTGAGAGAGTGAAGAGAGACTCCGAAAGTGTGGCTTCCGAGAAAGAGAGACG	3240
QY	3241	TTTAGCCGAAAGACCCGCTCAGAAAGCTGGAGCTGAGACATTTCTACACAGGAGCTTGAGA	3300
Db	3241	TTTAGCCGAAAGACCCGCTCAGAAAGCTGGAGCTGAGACATTTCTACACAGGAGCTTGAGA	3300
QY	3301	CTGACACAGGCTACTCCAGAGAGGCTTGAAGTTCTGAACTCAGCAGATGAGACCATG	3360
Db	3301	CTGACACAGGCTACTCCAGAGAGGCTTGAAGTTCTGAACTCAGCAGATGAGACCATG	3360
QY	3361	TGCGCCAGTGCAGGTTATCAAGCTCCAGAGCTGATGAGACAGGCGCTGGCCCTTGAGT	3420
Db	3361	TGCGCCAGTGCAGGTTATCAAGCTCCAGAGCTGATGAGACAGGCGCTGGCCCTTGAGT	3420
QY	3421	CATCCGAAACCTTGACAGACAGTGAACAAATGGAAAGCATCCCTTAGCAGATTACAGA	3480
Db	3421	CATCCGAAACCTTGACAGACAGTGAACAAATGGAAAGCATCCCTTAGCAGATTACAGA	3480
QY	3481	CTGCAGATGGGACACACAGATGAAACCATTGACAGCCAGGACAGTAAAGCATCTGCAG	3540
Db	3481	CTGCAGATGGGACACACAGATGAAACCATTGACAGCCAGGACAGTAAAGCATCTGCAG	3540
QY	3541	CTGTACGGCACTCACAGGTACAGAAAGAGGCGGCTACTGCTCAGAAAGAGAGCTT	3600
Db	3541	CTGTACGGCACTCACAGGTACAGAAAGAGGCGGCTACTGCTCAGAAAGAGAGCTT	3600
QY	3601	CGACACTTACCTTAATTAATGTTCCAGCCGAGAAACATATGGGAAAGAACACGAGAACAGT	3660
Db	3601	CGACACTTACCTTAATTAATGTTCCAGCCGAGAAACATATGGGAAAGAACACGAGAACAGT	3660
QY	3661	TTCTTGAACCTTACACACAGAGCTTACTGCTGCAGCCGCTGCGCTTGGCAAGCATG	3720
Db	3661	TTCTTGAACCTTACACACAGAGCTTACTGCTGCAGCCGCTGCGCTTGGCAAGCATG	3720
QY	3721	AGGTGGGTCAAGAGGGGTGAGTGAATGCTGGTAGTGAAGAAAGTCAAGAGAACAGG	3780
Db	3721	AGGTGGGTCAAGAGGGGTGAGTGAATGCTGGTAGTGAAGAAAGTCAAGAGAACAGG	3780
QY	3781	AGGCTTTTGTACTCTTGAGACCAACAGTCAAAAGGCTGTGATGTGACATATGACATG	3840
Db	3781	AGGCTTTTGTACTCTTGAGACCAACAGTCAAAAGGCTGTGATGTGACATATGACATG	3840
QY	3841	AAGTGATGGAGTGGCCGGGTGTCCAGAAAGAGAGTACTGAAGTGCAGACTTTAGCC	3900
Db	3841	AAGTGATGGAGTGGCCGGGTGTCCAGAAAGAGAGTACTGAAGTGCAGACTTTAGCC	3900
QY	3901	TGAGAGAGGAGAGATGAGAACTGACCTTGAAGAAAGAGAAAGGAGACAAAGCCAGAC	3960
Db	3901	TGAGAGAGGAGAGATGAGAACTGACCTTGAAGAAAGAGAAAGGAGACAAAGCCAGAC	3960
QY	3961	AAGTGAATGAAGAGGTGAGAGAGAAACACCGCTCTGACATGAAAGAACCTTACGGGAA	4020
Db	3961	AAGTGAATGAAGAGGTGAGAGAGAAACACCGCTCTGACATGAAAGAACCTTACGGGAA	4020
QY	4021	AGCCAGTCTTACACTTGACATGCTCCAGCTCAGAGAGGGGAGAGGACCTGGAGACCTTG	4080
Db	4021	AGCCAGTCTTACACTTGACATGCTCCAGCTCAGAGAGGGGAGAGGACCTGGAGACCTTG	4080
QY	4081	GAGGAAGCCCTTCTCTCCAGACCCAGACAAAGCAGGTGCTATGAGTTCAAGTTCAAA	4140
Db	4081	GAGGAAGCCCTTCTCTCCAGACCCAGACAAAGCAGGTGCTATGAGTTCAAGTTCAAA	4140
QY	4141	GCTTGAACACAAAGTCACTTAAACAGCACAACCTTGGAAAAAGTCAATGAAGAAAGGTTG	4200
Db	4141	GCTTGAACACAAAGTCACTTAAACAGCACAACCTTGGAAAAAGTCAATGAAGAAAGGTTG	4200
QY	4201	TGATTTCAAGACACAGGTGAAGTCCAGAGTGTGAGGTGACACTTATTCACAGCTGAGA	4260
Db	4201	TGATTTCAAGACACAGGTGAAGTCCAGAGTGTGAGGTGACACTTATTCACAGCTGAGA	4260
QY	4261	AGTCTCTTGCACACGGGTGGCCACTTGACCTTTTCAGATGACAGAGACACGATACCTTGG	4320

Db	4261	AGTCTCTGCAAGGGGTGGCCTACTGGA	CTTTGACATGCAAGAGACGGTACCTTGG	4320
Qy	4321	GGCCTGAGTCTCAGGAGAGATTCATCCCA	TATATGTAACCTCTGCTCTGAAAGCACCC	4380
Db	4321	GGCCTGAGTCTCAGGAGAGATTCATCCCA	TATATGTAACCTCTGCTCTGAAAGCACCC	4380
Qy	4381	TACATCTGACCTTACAGAGAAATTAAGCG	CAATGCCAGAGAGCGATCAGAGAGAGG	4440
Db	4381	TACATCTGACCTTACAGAGAAATTAAGCG	CAATGCCAGAGAGCGATCAGAGAGAGG	4440
Qy	4441	ACAAGCCAGATGTCGGTCCGATAGCTGAG	CGGAGAGAGAGATACAGCATGGAATAATCC	4500
Db	4441	ACAAGCCAGATGTCGGTCCGATAGCTGAG	CGGAGAGAGAGATACAGCATGGAATAATCC	4500
Qy	4501	TCAAGGCTGAACCTGAGATCTTGGAAC	CTTGAGAGTMAAGCAACAAGATTGTGCTGACG	4560
Db	4501	TCAAGGCTGAACCTGAGATCTTGGAAC	CTTGAGAGTMAAGCAACAAGATTGTGCTGACG	4560
Qy	4561	TCATTTCAAGACAGCCGTTTGACCAAGT	TGCGACGTACAGAAACAGCCCGGAACTCATGCTT	4620
Db	4561	TCATTTCAAGACAGCCGTTTGACCAAGT	TGCGACGTACAGAAACAGCCCGGAACTCATGCTT	4620
Qy	4621	ATGATTCACAGACCCAGGTTCTGCAATG	AGGCTTGCACAGAGAGGCGCCAAAGATGCT	4680
Db	4621	ATGATTCACAGACCCAGGTTCTGCAATG	AGGCTTGCACAGAGAGGCGCCAAAGATGCT	4680
Qy	4681	GGACAAAAATGAAAGATGCCCCAAGAT	AAAAACCCAGTCCGAGCCGAGAGAGACTTTCG	4740
Db	4681	GGACAAAAATGAAAGATGCCCCAAGAT	AAAAACCCAGTCCGAGCCGAGAGAGACTTTCG	4740
Qy	4741	AAATCTCGACCGCTTCTGAGAGCAATGG	GGCTTCAGCTTCGAGAAATGCTTGGCCGCTTTCGACT	4800
Db	4741	AAATCTCGACCGCTTCTGAGAGCAATGG	GGCTTCAGCTTCGAGAAATGCTTGGCCGCTTTCGACT	4800
Qy	4801	TGAAAAGGCGCGGTGTCAAAGTAAAGCA	TTGAGAAAGCGGCTCTCAACCCAAAGATCCAA	4860
Db	4801	TGAAAAGGCGCGGTGTCAAAGTAAAGCA	TTGAGAAAGCGGCTCTCAACCCAAAGATCCAA	4860
Qy	4861	AGGAGCAATGCTGCTGATGAGCGCTCA	CACTTCAAGGTTAGCCGAGGACAGGCGCAGTGCCT	4920
Db	4861	AGGAGCAATGCTGCTGATGAGCGCTCA	CACTTCAAGGTTAGCCGAGGACAGGCGCAGTGCCT	4920
Qy	4921	CTGGAATCTTAAACCAAGATTCCTCCAG	ACCAACGATCCGAGAGAGGCG	4980
Db	4921	CTGGAATCTTAAACCAAGATTCCTCCAG	ACCAACGATCCGAGAGAGGCG	4980
Qy	4981	ATCCCCCAAAAGTTGAGAGTCCAGAGAA	AGAAATGCTTACCAAGTCACTCAAGAGAGAA	5040
Db	4981	ATCCCCCAAAAGTTGAGAGTCCAGAGAA	AGAAATGCTTACCAAGTCACTCAAGAGAGAA	5040
Qy	5041	AGGCCACAGGAGAGAGACTGACAGAGC	CAAAAGGAGAGCTTGACAGATCTTAAATG	5100
Db	5041	AGGCCACAGGAGAGAGACTGACAGAGC	CAAAAGGAGAGCTTGACAGATCTTAAATG	5100
Qy	5101	TTAGTGTCTATTGTATCTTGTAAACCA	GAATGTGAATAAAGTCAAGAAACAAGATG	5160
Db	5101	TTAGTGTCTATTGTATCTTGTAAACCA	GAATGTGAATAAAGTCAAGAAACAAGATG	5160
Qy	5161	CTGCGTTGGGACCTTGAGACCAAGATTT	CAGAGCCAGATTCAGAGAGAGGCGG	5220
Db	5161	CTGCGTTGGGACCTTGAGACCAAGATTT	CAGAGCCAGATTCAGAGAGAGGCGG	5220
Qy	5221	TCCAAATGATTTCCACCAGTAGAGACACC	CGACAATTTGAGAGCTTCAATCGGAGACTGAG	5280
Db	5221	TCCAAATGATTTCCACCAGTAGAGACACC	CGACAATTTGAGAGCTTCAATCGGAGACTGAG	5280
Qy	5281	GCCAGCTAAACATTTCTCTGTTCAAGAT	CTGCTTGATTTGCCCTTATGATGCCGTCGCTG	5340
Db	5281	GCCAGCTAAACATTTCTCTGTTCAAGAT	CTGCTTGATTTGCCCTTATGATGCCGTCGCTG	5340
Qy	5341	TATTTCTAAACATTTCTCTGTTCAAGAT	CTGCTTGATTTGCCCTTATGATGCCGTCGCTG	5400
Db	5341	TATTTCTAAACATTTCTCTGTTCAAGAT	CTGCTTGATTTGCCCTTATGATGCCGTCGCTG	5400

QY	601	GGGAGAAAGCTCTGTGGAGCTGGAGACCA	CCAGAGGCCAGTGTGGAGACTGCGCTG	660
Db	601	GGGGAGAAAGCTCTGTGGAGCTGGAGACCA	CCAGAGGCCAGTGTGGAGACTGCGCTG	660
QY	661	GAGAGTCAGCATCCAAAGAAAGTAGTGAAGCA	ATCCACACAGAGAAGCAAGAGCACCC	720
Db	661	GAGAGTCAGCATCCAAAGAAAGTAGTGAAGCA	ATCCACAGAGAAGCAAGAGCACCC	720
QY	721	TGAAGCAGAAACAGAGCAGACAGAAATCCCC	CTTCAAGCCGAAATCTGATCAAGCGGCTG	780
Db	721	TGAAGCAGAAACAGAGCAGACAGAAATCCCC	CTTCAAGCCGAAATCTGATCAAGCGGCTG	780
QY	781	AGGAAGAAAGCCAAAGATGAAGAGAGAA	AAACAGAGAAAGAGCCCAACAGTCTCCAG	840
Db	781	AGGAAGAAAGCCAAAGATGAAGAGAGAA	AAACAGAGAAAGAGCCCAACAGTCTCCAG	840
QY	841	AAATCCCGAGAGCCCAAGTCAACAGTAGACA	ATCTTCTTCAAGAAAGTTCTTCACTC	900
Db	841	AAATCCCGAGAGCCCAAGTCAACAGTAGACA	ATCTTCTTCAAGAAAGTTCTTCACTC	900
QY	901	ACGGTTGGGCGCGCTGGCGCAAGAGAC	CAGCTTCAAGAAATCAAAAAGAGATGATCGG	960
Db	901	ACGGTTGGGCGCGCTGGCGCAAGAGAC	CAGCTTCAAGAAATCAAAAAGAGATGATCGG	960
QY	961	AAACTGCCGAGAGAGAAAGAGCAGAGCGC	CAGAAAAGTGAACGAGAGAAAGAGAA	1020
Db	961	AAACTGCCGAGAGAGAAAGAGCAGAGCGC	CAGAAAAGTGAACGAGAGAAAGAGAA	1020
QY	1021	AGACAGAGCCAGCCTCGAGAGAGAGAGAGC	CGGCAABAACACAGCCAGAGCTGT	1080
Db	1021	AGACAGAGCCAGCCTCGAGAGAGAGAGAGC	CGGCAABAACACAGCCAGAGCTGT	1080
QY	1081	CAGCAGACTTACGAAGAAGGTGAGCTGTG	TGAAAGACAGAGTTGGTGA	1140
Db	1081	CAGCAGACTTACGAAGAAGGTGAGCTGTG	TGAAAGACAGAGTTGGTGA	1140
QY	1141	CGTCAAGAGAGAGTGTGTCTCTTTGGCA	ACGAAAGTTTATGAGAAATGGAAGCC	1200
Db	1141	CGTCAAGAGAGAGTGTGTCTCTTTGGCA	ACGAAAGTTTATGAGAAATGGAAGCC	1200
QY	1201	ACCAAGAAAGTTGTGACAGAGGTCCA	CGTGAAGCAACGGTGAAGAAACAGAGAGAG	1260
Db	1201	ACCAAGAAAGTTGTGACAGAGGTCCA	CGTGAAGCAACGGTGAAGAAACAGAGAGAG	1260
QY	1261	GAGGAGAGAGAGGCTGAGAGGGGCGCTG	TGTAAGAAACAGAGATCTTGCCCC	1320
Db	1261	GAGGAGAGAGAGGCTGAGAGGGGCGCTG	TGTAAGAAACAGAGATCTTGCCCC	1320
QY	1321	CTGAGAAA	CTGGCTGAGCCCCCAGAGAGTCCCC	1380
Db	1321	CTGAGAAA	CTGGCTGAGCCCCCAGAGAGTCCCC	1380
QY	1381	TGAAGAGCAGAGATGTGTCTCTTGAAGAGA	CCACACTCACTGACAGACCTAAATC	1440
Db	1381	TGAAGAGCAGAGATGTGTCTCTTGAAGAGA	CCACACTCACTGACAGACCTAAATC	1440
QY	1441	CTGAAGAGAAACCGCTGCCCAAC	CAACCAAGAGCATTTGTCACTGAGTGGAGATGCTGT	1500
Db	1441	CTGAAGAGAAACCGCTGCCCAAC	CAACCAAGAGCATTTGTCACTGAGTGGAGATGCTGT	1500
QY	1501	CCTCTCAGAGAAAGATCAAGGTACAGGGAA	ATCCCTTTGAAGAAATCTTCACTAGCTCAG	1560
Db	1501	CCTCTCAGAGAAAGATCAAGGTACAGGGAA	ATCCCTTTGAAGAAATCTTCACTAGCTCAG	1560
QY	1561	GCTTAAAGAACTGTCTGGGAAAGAGCA	AGAGGGGAAACAGAGAGTGGGGGAGACAGAG	1620
Db	1561	GCTTAAAGAACTGTCTGGGAAAGAGCA	AGAGGGGAAACAGAGAGTGGGGGAGACAGAG	1620
QY	1621	AGCCTGAGAAATACCAACACTTCAKAC	CGAAATCCCAAGAGTGTCTGATGAGCAGAAAG	1680
Db	1621	AGCCTGAGAAATACCAACACTTCAKAC	CGAAATCCCAAGAGTGTCTGATGAGCAGAAAG	1680
QY	1681	GAGAGAGCTCTGCGTGTCTCCCGAGAG	AGCCTTGAAGACACGTGTCTTGAAGAAAGGC	1740

[illegible]

Db 2761 TGCCACCTGTTGAAGAGGTCACTGAAAAGACATCATTTGCAGAGAAACTCCTCTGTGCTCA 2820
Qy 2821 CCCAGAGCTTACAGAGGTTAAAGATGCCCATGACGACATGGTCCAGTGAAATGAGATT 2880
Db 2821 CCCAGAGCTTACAGAGGTTAAAGATGCCCATGACGACATGGTCCAGTGAAATGAGATT 2880
Qy 2881 TCACCTCAGAAAGCTGTGACACGCCACAGAGACCTCAGAGGCTCTCGTACTGAAAGATT 2940
Db 2881 TCACCTCAGAAAGCTGTGACACGCCACAGAGACCTCAGAGGCTCTCGTACTGAAAGATT 2940
Qy 2941 CCGAAGCATCGGGGGCCGAAGAGAGCCACAGACATGGTGTCCGCACTGACTG 3000
Db 2941 CCGAAGCATCGGGGGCCGAAGAGAGCCACAGACATGGTGTCCGCACTGACTG 3000
Qy 3001 ACTCCCCAGACACACAGAGAGAACCCACCCAGTTCAGGAGGTAGAGTGTGTGTAG 3060
Db 3001 ACTCCCCAGACACACAGAGAGAACCCACCCAGTTCAGGAGGTAGAGTGTGTGTAG 3060
Qy 3061 ATACAGAAGAGAGGAGCGCCAGACGACGAGGCATCTCCAAAGCCGTTCGACACAAGTGA 3120
Db 3061 ATACAGAAGAGAGGAGCGCCAGACGACGAGGCATCTCCAAAGCCGTTCGACACAAGTGA 3120
Qy 3121 AAGAGAGTCCCAAGTGCCTGCAACCCAGACTGTGCAAGAGAACCGGGTCAAAAAGCACTGG 3180
Db 3121 AAGAGAGTCCCAAGTGCCTGCAACCCAGACTGTGCAAGAGAACCGGGTCAAAAAGCACTGG 3180
Qy 3181 AGAAGTTGAGGAGGTAGAGGAGACTCCGAAGTGTCTGGTTCGGAGAAAGAGAGACG 3240
Db 3181 AGAAGTTGAGGAGGTAGAGGAGACTCCGAAGTGTCTGGTTCGGAGAAAGAGAGACG 3240
Qy 3241 TTATGCCGAAAGGACCCGTGCGAGAGCTGGAGCTGAGCATCTTGCAAGGGCTCTGAGA 3300
Db 3241 TTATGCCGAAAGGACCCGTGCGAGAGCTGGAGCTGAGCATCTTGCAAGGGCTCTGAGA 3300
Qy 3301 CTGAGACGGCTACTCCAGAGAGCCTTGAAGTTCCTGAAGTCAAGGAGATGTAGACCATG 3360
Db 3301 CTGAGACGGCTACTCCAGAGAGCCTTGAAGTTCCTGAAGTCAAGGAGATGTAGACCATG 3360
Qy 3361 TCGCCACGTCGCGCTTATCAAGCTCCAGCAGCTGATGGAAACAGGCGCTGGCCCTGAGT 3420
Db 3361 TCGCCACGTCGCGCTTATCAAGCTCCAGCAGCTGATGGAAACAGGCGCTGGCCCTGAGT 3420
Qy 3421 CATCCGAAACCTTGACAGACAGTGAGACAAATGGAAGCATCCCTTAGCAGATTTCAGACA 3480
Db 3421 CATCCGAAACCTTGACAGACAGTGAGACAAATGGAAGCATCCCTTAGCAGATTTCAGACA 3480
Qy 3481 CTGCAGATGGGACACAGCAAGATCAAACTTGTACAGCAGGACAGTAAAGCCACTGCAG 3540
Db 3481 CTGCAGATGGGACACAGCAAGATCAAACTTGTACAGCAGGACAGTAAAGCCACTGCAG 3540
Qy 3541 CTGTCAGGCACTCAGAGTCAAGAGAGAGGCGGCTACTGCTCAGAAAAGAGAGCCTT 3600
Db 3541 CTGTCAGGCACTCAGAGTCAAGAGAGAGGCGGCTACTGCTCAGAAAAGAGAGCCTT 3600
Qy 3601 CGACACTACCTAATATGTTCCAGCCAGGAAGAACATGGGGAAGAACCCAGGAAGAGATG 3660
Db 3601 CGACACTACCTAATATGTTCCAGCCAGGAAGAACATGGGGAAGAACCCAGGAAGAGATG 3660
Qy 3661 TTCTTTGAACCTTACACAGCAGAGCTTACTGCTGAGCCGCTGCGGCTTCTGCAAGACTG 3720
Db 3661 TTCTTTGAACCTTACACAGCAGAGCTTACTGCTGAGCCGCTGCGGCTTCTGCAAGACTG 3720
Qy 3721 AGTGGGTCAAGAGGGTGAGGTTGACTGGTTGGATGGAGAAAAAGTCAAAAGAAACAGG 3780
Db 3721 AGTGGGTCAAGAGGGTGAGGTTGACTGGTTGGATGGAGAAAAAGTCAAAAGAAACAGG 3780
Qy 3781 AGTGGTTGTACACTTGGACCCCAAGTCAAAAGGCTGTGATGTGACATATGACAGTG 3840
Db 3781 AGTGGTTGTACACTTGGACCCCAAGTCAAAAGGCTGTGATGTGACATATGACAGTG 3840
Qy 3841 AAGTGATGGGAGTGGCGGGTGTGAGGAAAGGAGAGTACTGAAAGTGCAGAGTCTTAGCC 3900
Db 3841 AAGTGATGGGAGTGGCGGGTGTGAGGAAAGGAGAGTACTGAAAGTGCAGAGTCTTAGCC 3900

Qy 3901 TGGAGGAGGAGAGATGGAAACTGACCTTGAAAAGGAGAAAAGGAGACAAAAGCCAGAGC 3960
Db 3901 TGGAGGAGGAGAGATGGAAACTGACCTTGAAAAGGAGAAAAGGAGACAAAAGCCAGAGC 3960
Qy 3961 AAGTGTAGTGAAGAGGTGAGCAGGAAACACGCCCTCTCTGAGCATGGAAGAACCTTACCGGA 4020
Db 3961 AAGTGTAGTGAAGAGGTGAGCAGGAAACACGCCCTCTCTGAGCATGGAAGAACCTTACCGGA 4020
Qy 4021 AGCCAGTCTTGACACTTGACATGCCAGCTCAGAGAGGGGGAAGGCACTGGGAAGCCTTG 4080
Db 4021 AGCCAGTCTTGACACTTGACATGCCAGCTCAGAGAGGGGGAAGGCACTGGGAAGCCTTG 4080
Qy 4081 GAGGAAGCCCTTCTCTCCAGACCAAGAAGAGAGGTTGCATAGAGGTTCAAGTTCAAA 4140
Db 4081 GAGGAAGCCCTTCTCTCCAGACCAAGAAGAGAGGTTGCATAGAGGTTCAAGTTCAAA 4140
Qy 4141 GCCTGGACACAAAGTCACTCAAAACAGCAGCAAGCTGTGGAAAAGGTCATAGAACGGTTG 4200
Db 4141 GCCTGGACACAAAGTCACTCAAAACAGCAGCAAGCTGTGGAAAAGGTCATAGAACGGTTG 4200
Qy 4201 TGATTTTCAGAGACAGGTGAAAGTCCAGAGTGTGTAGGTGCACACTTTATTACCAGCTGAGA 4260
Db 4201 TGATTTTCAGAGACAGGTGAAAGTCCAGAGTGTGTAGGTGCACACTTTATTACCAGCTGAGA 4260
Qy 4261 AGTCTCTTGCAAAGGGTGGCCACTGGACTCTTTCAGCATGCAGAGGACACGGTACCCCTGG 4320
Db 4261 AGTCTCTTGCAAAGGGTGGCCACTGGACTCTTTCAGCATGCAGAGGACACGGTACCCCTGG 4320
Qy 4321 GGCTGTAGTCTCAGGAGAGATCCATCCCAATCATAGTAACTCTCTGCTTGAAGACACC 4380
Db 4321 GGCTGTAGTCTCAGGAGAGATCCATCCCAATCATAGTAACTCTCTGCTTGAAGACACC 4380
Qy 4381 TACATCTCTGACCTTACAGGAGAAATTAAGCGCATCCAGAGAGAGCGATCAGAGGAAGAGG 4440
Db 4381 TACATCTCTGACCTTACAGGAGAAATTAAGCGCATCCAGAGAGAGCGATCAGAGGAAGAGG 4440
Qy 4441 ACAAGCCAGATGCTGTCTGTATGCTGACGGCAAGGAGAGTACAGCAATCGAAAAAGTCC 4500
Db 4441 ACAAGCCAGATGCTGTCTGTATGCTGACGGCAAGGAGAGTACAGCAATCGAAAAAGTCC 4500
Qy 4501 TCAAGCTGAACTGAGATCTCTGAACTTGAAGATTAAGAGCAACAGATTTGTCTGAACG 4560
Db 4501 TCAAGCTGAACTGAGATCTCTGAACTTGAAGATTAAGAGCAACAGATTTGTCTGAACG 4560
Qy 4561 TCATTTCAGACAGCGGTTGACCACTGCGAGTTCGAGTACAGAAACAGCCCGGAACTCATGCTT 4620
Db 4561 TCATTTCAGACAGCGGTTGACCACTGCGAGTTCGAGTACAGAAACAGCCCGGAACTCATGCTT 4620
Qy 4621 ATGATTCACAGACCCAGGTTCTGTCATGAGGTTGACAGAGGAGCCCAACAGATGCT 4680
Db 4621 ATGATTCACAGACCCAGGTTCTGTCATGAGGTTGACAGAGGAGCCCAACAGATGCT 4680
Qy 4681 GGACAAAATGAAAGATGCCAAGATGAAACACCCAGTTCGCGAGCCGAGAGAGCACTTGC 4740
Db 4681 GGACAAAATGAAAGATGCCAAGATGAAACACCCAGTTCGCGAGCCGAGAGAGCACTTGC 4740
Qy 4741 AAGTCTGACCGTTCTGGAGGATGGCTCAGCTCGGAAATGCTTGGCGGCTTGCAGT 4800
Db 4741 AAGTCTGACCGTTCTGGAGGATGGCTCAGCTCGGAAATGCTTGGCGGCTTGCAGT 4800
Qy 4801 TGAAGCGCCGGTGTCAAAGTAAAGCATTTGAGAAAGTGCCTCTCAACCCCAAGATCCAAA 4860
Db 4801 TGAAGCGCCGGTGTCAAAGTAAAGCATTTGAGAAAGTGCCTCTCAACCCCAAGATCCAAA 4860
Qy 4861 AGGAGCATGCTGTATGGCCCTCAGCTCAAAAGTGTAGCCAGGAGAGGCGCAGTGCCT 4920
Db 4861 AGGAGCATGCTGTATGGCCCTCAGCTCAAAAGTGTAGCCAGGAGAGGCGCAGTGCCT 4920
Qy 4921 CTGGAACCTTAACCAAGNATCCCGAGACCAACCGAACCAAGCTTAACCGAGAGGGCG 4980
Db 4921 CTGGAACCTTAACCAAGNATCCCGAGACCAACCGAACCAAGCTTAACCGAGAGGGCG 4980

Oy	4981	ATCCCCAAAGTTGAGGCTCAAGAAAGAAATGTCTACAAAGTCAGTCAAGAGAA	5040
Db	4981	ATCCCCAAAGTTGAGGCTCAAGAAAGAAATGTCTACAAAGTCAGTCAAGAGAA	5040
Oy	5041	AGGCCAGCAGAAAGAGGAGCTGCAGAGAGCAAAAGGAGACTGGCAGAACTTAAGATG	5100
Db	5041	AGGCCAGCAGAAAGAGGAGCTGCAGAGAGCAAAAGGAGACTGGCAGAACTTAAGATG	5100
Oy	5101	TTAGTTCCTCATGTGTACATCTGTAAACCAAGATGTGAAAACAAATCAAGAACAAATG	5160
Db	5101	TTAGTTCCTCATGTGTACATCTGTAAACCAAGATGTGAAAACAAATCAAGAACAAATG	5160
Oy	5161	CTGCTGTGGGACCTTGAGACCAAGATTTTCAGAGCCCATGAGATCCAGAGAGCAGGGCCG	5220
Db	5161	CTGCTGTGGGACCTTGAGACCAAGATTTTCAGAGCCCATGAGATCCAGAGAGCAGGGCCG	5220
Oy	5221	TCCATGATTTTCCACCAGTAGAGACACCCGCAATTCGAGAGCTTCATCGGAGACTAGA	5280
Db	5221	TCCATGATTTTCCACCAGTAGAGACACCCGCAATTCGAGAGCTTCATCGGAGACTAGA	5280
Oy	5281	GCCAGCTAACATTTTCCTCGTTTCAAGACTGCTGATTGGCCCTTGATGTCGTCGCTG	5340
Db	5281	GCCAGCTAACATTTTCCTCGTTTCAAGACTGCTGATTGGCCCTTGATGTCGTCGCTG	5340
Oy	5341	TATTTCTAACATTTTCTCTGTTTCAAGACTGCCCTTTGATTTGGCCCTTGATGTCGTCGCTG	5400
Db	5341	TATTTCTAACATTTTCTCTGTTTCAAGACTGCCCTTTGATTTGGCCCTTGATGTCGTCGCTG	5400
Oy	5401	TATTTGGATTTAAGTCCGCGCTTCACACTCGAACAATTCCTGSCATACCTAGTTC	5460
Db	5401	TATTTGGATTTAAGTCCGCGCTTCACACTCGAACAATTCCTGSCATACCTAGTTC	5460
Oy	5461	ACTTCTCAACTGAGAGACTCTCCTTTATGATTTATGATGTTTTATGTAGTCTCC	5520
Db	5461	ACTTCTCAACTGAGAGACTCTCCTTTATGATTTATGATGTTTTATGTAGTCTCC	5520
Oy	5521	TCTCTTACCTATGTATATTTTTTTTCTAACGTTTAGACAACGCTTTGATTTATAGAA	5580
Db	5521	TCTCTTACCTATGTATATTTTTTTTCTAACGTTTAGACAACGCTTTGATTTATAGAA	5580
Oy	5581	TATATTAACGGGTGACAGCCATAGCAGCGTTTGAAAAGCTCCAGGCCCTCAACTTAAC	5640
Db	5581	TATATTAACGGGTGACAGCCATAGCAGCGTTTGAAAAGCTCCAGGCCCTCAACTTAAC	5640
Oy	5641	TGCAGCAACAGATTAACATTCCTGGCAAGAGAGACAGTCTTTTAAAGTTATCTAT	5700
Db	5641	TGCAGCAACAGATTAACATTCCTGGCAAGAGAGACAGTCTTTTAAAGTTATCTAT	5700
Oy	5701	GCTTAGATCTGTGGGCTTCTAGTCTCTGAAAGTGTTGTTTCCATAGCAAGGGACT	5760
Db	5701	GCTTAGATCTGTGGGCTTCTAGTCTCTGAAAGTGTTGTTTCCATAGCAAGGGACT	5760
Oy	5761	CAGAAATTAACACCCCAATTTGAAACATCAAGATGTCCCAATATTAACATGATTTTTC	5820
Db	5761	CAGAAATTAACACCCCAATTTGAAACATCAAGATGTCCCAATATTAACATGATTTTTC	5820
Oy	5821	CCCCCTTTTCTATCCAGTCCAGGTGGAAAAGTCTCCTCTGTGCAGATTAAGC	5880
Db	5821	CCCCCTTTTCTATCCAGTCCAGGTGGAAAAGTCTCCTCTGTGCAGATTAAGC	5880
Oy	5881	CTGTCTTTAATGATATGACAAATGAGTGTGCTTAAGGCCATGAGATGTTTCTTAATGC	5940
Db	5881	CTGTCTTTAATGATATGACAAATGAGTGTGCTTAAGGCCATGAGATGTTTCTTAATGC	5940
Oy	5941	AGAAAGAACTGTGTAGAGTTTTTAATGTATCTCTTAAGTGTGACCAATTCATAT	6000
Db	5941	AGAAAGAACTGTGTAGAGTTTTTAATGTATCTCTTAAGTGTGACCAATTCATAT	6000
Oy	6001	GCAGATCGAAGGAGTCTGTCTTTTCAAGATGATATTTGATAGATACCTGAGATTTGTC	6060
Db	6001	GCAGATCGAAGGAGTCTGTCTTTTCAAGATGATATTTGATAGATACCTGAGATTTGTC	6060
Oy	6061	TGTGTATATCTGTGCCCTTCTTTAAGAACATGTTGCATATGTTCCTTGGATTAAT	6120

Db 6061 TGTGTTATATCTGTGGCCCTCTTTAAGAACAGTGTGATTATGTCTCTTGATATAAT 6120

Qy 6121 TGTGATTTGACAACTGATTTAAATAAACATATTGTGACTAC 6160

Db 6121 TGTGATTTGACAACTGATTTAAATAAACATATTGTGACTAC 6160

```

RESULT 3
US-09-902-432-3
: Sequence 3, Application US/09902432
: GENERAL INFORMATION:
:
: APPLICANT: Irwin H. Gelman
: APPLICANT: Susan G. Jaken
: TITLE OF INVENTION: Tumor Suppressor Gene
: FILE REFERENCE: A30558-A-FWC-A 070156.05
: CURRENT APPLICATION NUMBER: US/09/902.432
: CURRENT FILING DATE: 2002-04-08
: PRIOR APPLICATION NUMBER: 08/978,277
: PRIOR FILING DATE: 1997-11-25
: PRIOR APPLICATION NUMBER: 08/665,401
: PRIOR FILING DATE: 1996-06-18
: PRIOR APPLICATION NUMBER: 08/635,121
: PRIOR FILING DATE: 1996-04-19
: NUMBER OF SEQ ID NOS: 35
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 3
: LENGTH: 6160
: TYPE: DNA
: ORGANISM: Rattus norvegicus
US-09-902-432-3

```

Query Match	100.0%	Score 6160;	DB 34;	Length 6160;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 6160;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	1	CGGGCCGGGAGTAGAAGACCACTGACCCATGCGGCGAGGACAGTTCCACCGACAGC	60
Db	1	CGGGCCGGGAGTAGAAGACCACTGACCCATGCGGCGAGGACAGTTCCACCGACAGC	60
Qy	61	GGAGCCCGGAGCAGCCGAGCGGGAGCGACAACGCCGAGCGAGCTGTGTCTAGTGGCCATG	120
Db	61	GGAGCCCGGAGCAGCCGAGCGGGAGCGACAACGCCGAGCGAGCTGTGTCTAGTGGCCATG	120
Qy	121	GGCCCGCAGCTGAAAGCCTCGGGAGCACTGGAGACCCCCCGGACGGGACCCCGCCACA	180
Db	121	GGCCCGCAGCTGAAAGCCTCGGGAGCACTGGAGACCCCCCGGACGGGACCCCGCCACA	180
Qy	181	AGCTCCCAACAAGAAATGGCAGCTGTCTTCTGTCAACGGCGTAAAGCTGACAAAGAGATG	240
Db	181	AGCTCCCAACAAGAAATGGCAGCTGTCTTCTGTCAACGGCGTAAAGCTGACAAAGAGATG	240
Qy	241	TCCATGTCCAAAGGAAAAACAGAGGGGCGAGAGAAAGTCGTTGATGAGAGATGTG	300
Db	241	TCCATGTCCAAAGGAAAAACAGAGGGGCGAGAGAAAGTCGTTGATGAGAGATGTG	300
Qy	301	GACACGAGCTCAGAACATGTGAGAAAAAGACCGAGTTGAAGAAATGGCGGCAACT	360
Db	301	GACACGAGCTCAGAACATGTGAGAAAAAGACCGAGTTGAAGAAATGGCGGCAACT	360
Qy	361	CCACAGCTGTGGAAGATTCACAAAGAGTGGCGAGAGGACATCATCAAAATTAATGAAAC	420
Db	361	CCACAGCTGTGGAAGATTCACAAAGAGTGGCGAGAGGACATCATCAAAATTAATGAAAC	420
Qy	421	AGATCCCGCTTCAGAAAACAATGGAGAAATGTAACAGCCCTCGAGATCCAGGCTA	480
Db	421	AGATCCCGCTTCAGAAAACAATGGAGAAATGTAACAGCCCTCGAGATCCAGGCTA	480
Qy	481	ATGATGTGGCTTCAGAAAGATATTAATTTGTGGTTTAAATTCACCGTGAAGAGG	540
Db	481	ATGATGTGGCTTCAGAAAGATATTAATTTGTGGTTTAAATTCACCGTGAAGAGG	540
Qy	541	ATTAATAATGAAAAATCAGATATCTGTCCAATCACTGTCAACAGAGATGAAGCGAAG	6000

Db 541 ATAAAAATGAAAGTCAGATACTCTCCAACTACTCTCAAGAAAGGATCAAGCGCAG 600
Qy 601 GGGCAGAAAGCTCTGTGCGAGCTCGAGACCAACAGGAGCCAGTGTGAGACTCCCTCG 660
Db 601 GGGCAGAAAGCTCTGTGCGAGCTCGAGACCAACAGGAGCCAGTGTGAGACTCCCTCG 660
Qy 661 GAGAGTCAGCATCCAAAGAAAGTCAGTGAAGCAATCCACAGAGAAAGCAAGGCAACC 720
Db 661 GAGAGTCAGCATCCAAAGAAAGTCAGTGAAGCAATCCACAGAGAAAGCAAGGCAACC 720
Qy 721 TGAAGCAAGAACAGAGCAGACAGAAATCCCCCTTCAAGCCGAATCTGTCAAGCGCTG 780
Db 721 TGAAGCAAGAACAGAGCAGACAGAAATCCCCCTTCAAGCCGAATCTGTCAAGCGCTG 780
Qy 781 AGAAGAACCAAGATGAAGGAGAAAGAAACAAAGAGAAAGAGCCCAACCAAGTCCCCAG 840
Db 781 AGAAGAACCAAGATGAAGGAGAAAGAAACAAAGAGAAAGAGCCCAACCAAGTCCCCAG 840
Qy 841 AATCCCGAGCAGCCAGTCAACAGTCAAGCAGCAACATCTTCTCAAGAGTCTTCACTC 900
Db 841 AATCCCGAGCAGCCAGTCAACAGTCAAGCAGCAACATCTTCTCAAGAGTCTTCACTC 900
Qy 901 ACGTTGGCGCGGTGGCGCAAGAACCCAGCTTCAAGAAATCAAGAGAGATGATCTGG 960
Db 901 ACGTTGGCGCGGTGGCGCAAGAACCCAGCTTCAAGAAATCAAGAGAGATGATCTGG 960
Qy 961 AACTGCCGAGAGAGAAAGAGCAAGGAGCAAAAAGTAGACGAGGAGAAAGAAAGAAA 1020
Db 961 AACTGCCGAGAGAGAAAGAGCAAGGAGCAAAAAGTAGACGAGGAGAAAGAAAGAAA 1020
Qy 1021 AGACAGAGCAGCCTCGGAGGAGCAGAGCCGAGAGAAAGACACAGACAGGCCAGTTGT 1080
Db 1021 AGACAGAGCAGCCTCGGAGGAGCAGAGCCGAGAGAAAGACACAGACAGGCCAGTTGT 1080
Qy 1081 CAGCAGACTACGAAAGTGGAGCTGCTTTGGAAACCGAGTTGGTGTGACTGGAGGAT 1140
Db 1081 CAGCAGACTACGAAAGTGGAGCTGCTTTGGAAACCGAGTTGGTGTGACTGGAGGAT 1140
Qy 1141 CGTCAGAGGAGAGTGTGCTCTTTGGCAACGGAAGTGTGATGAGAAATGGAAGCCC 1200
Db 1141 CGTCAGAGGAGAGTGTGCTCTTTGGCAACGGAAGTGTGATGAGAAATGGAAGCCC 1200
Qy 1201 ACCAAGAGTGTGACAGAGTCCAGTGAACACCGTGGAGAGACAGAGAGAGCAGG 1260
Db 1201 ACCAAGAGTGTGACAGAGTCCAGTGAACACCGTGGAGAGACAGAGAGAGCAGG 1260
Qy 1261 GAGGAGGAGAGAGGCTGAAGGGGCGTGTGTGAGAGGACAGAGAGATCTTGGCCC 1320
Db 1261 GAGGAGGAGAGAGGCTGAAGGGGCGTGTGTGAGAGGACAGAGAGATCTTGGCCC 1320
Qy 1321 CTGAGAACTGGCTGAGCCCGAGAGGTCCTCCAGGAAGCTGAGCCTGCTGAGGAGCTGA 1380
Db 1321 CTGAGAACTGGCTGAGCCCGAGAGGTCCTCCAGGAAGCTGAGCCTGCTGAGGAGCTGA 1380
Qy 1381 TGAAGCAGACAGAGATGTGTCTGAGAGAGACACACTCAACTGACAGACCTAAGTC 1440
Db 1381 TGAAGCAGACAGAGATGTGTCTGAGAGAGACACACTCAACTGACAGACCTAAGTC 1440
Qy 1441 CTGAAGAGAGAGAGCTGCCCAACACCGAGAGGATGTGTGAGGAGGAGAGATCTGT 1500
Db 1441 CTGAAGAGAGAGAGCTGCCCAACACCGAGAGGATGTGTGAGGAGGAGAGATCTGT 1500
Qy 1501 CCTCTCAGGAAAGAACTCAAGGTACAGGAAAGTCCCTTGAAGAACTCTTCAAGTGTCTAG 1560
Db 1501 CCTCTCAGGAAAGAACTCAAGGTACAGGAAAGTCCCTTGAAGAACTCTTCAAGTGTCTAG 1560
Qy 1561 GCTTAAAGAGAGTGTCTGAGAGAGAGCAGAGAGGAGAAACAGAGAGGTGGGGAGACGAAG 1620
Db 1561 GCTTAAAGAGAGTGTCTGAGAGAGAGCAGAGAGGAGAAACAGAGAGGTGGGGAGACGAAG 1620
Qy 1621 AGCTGAGAGAAACCAACATTCACACCAATCCCGAGAGTGTGATGAGCAGAGG 1680

Db 1621 AGCTGAGAGAAACCAACATTCACACCAATCCCGAGAGAGTGTGATGAGCAGAGG 1680
Qy 1681 GAGAGAGCTCTGCTGCTCCCCGAGGAGCCTGAGGAGACCACTGTCTGAGAGAAAGGGC 1740
Db 1681 GAGAGAGCTCTGCTGCTCCCCGAGGAGCCTGAGGAGACCACTGTCTGAGAGAAAGGGC 1740
Qy 1741 CGCTGGAAGCAGCCAGGATGGGGAAGCTGAGGAAGAACTACTTCCGATGGGAGAGA 1800
Db 1741 CGCTGGAAGCAGCCAGGATGGGGAAGCTGAGGAAGAACTACTTCCGATGGGAGAGA 1800
Qy 1801 AGAGAGAGGAGTCACTCTCTGGGCACTCTTCAAAAAGATGGTGACACCCAAAGAACGG 1860
Db 1801 AGAGAGAGGAGTCACTCTCTGGGCACTCTTCAAAAAGATGGTGACACCCAAAGAACGG 1860
Qy 1861 TCCGAAGACCTCTGAGAGTGAACAAGGAGAGAGCTGGAGAAAGGTCAAGAGCGCCACCT 1920
Db 1861 TCCGAAGACCTCTGAGAGTGAACAAGGAGAGAGCTGGAGAAAGGTCAAGAGCGCCACCT 1920
Qy 1921 TGTCTCTCCACTGATAGCAGTGTCAAGATGCAAGATGAACTCAAACTGTTGGTGAGG 1980
Db 1921 TGTCTCTCCACTGATAGCAGTGTCAAGATGCAAGATGAACTCAAACTGTTGGTGAGG 1980
Qy 1981 AACAAAAGCCAGAGGAAACAAAGCGTAGGTGGATACCTCAGTGTCTTGGGAAAGCACTGA 2040
Db 1981 AACAAAAGCCAGAGGAAACAAAGCGTAGGTGGATACCTCAGTGTCTTGGGAAAGCACTGA 2040
Qy 2041 TTTGTCTCGGATCATCAAGAAAGAGAGCAAGGAGGATCTCTTTCAGATGATGAAGGAG 2100
Db 2041 TTTGTCTCGGATCATCAAGAAAGAGAGCAAGGAGGATCTCTTTCAGATGATGAAGGAG 2100
Qy 2101 GGCACAGGACACTGGGAGGGGACAGTCAACAGCAGAGAGGAGCCAGCAAGAAAG 2160
Db 2101 GGCACAGGACACTGGGAGGGGACAGTCAACAGCAGAGAGGAGCCAGCAAGAAAG 2160
Qy 2161 CCGGAAACAGACGCTGTCTTCCAGCAGCCAGGAGCAGGACCAAGCGCAGGAGGTTTCT 2220
Db 2161 CCGGAAACAGACGCTGTCTTCCAGCAGCCAGGAGCAGGACCAAGCGCAGGAGGTTTCT 2220
Qy 2221 CACCCGAGCCAGCGGAGCCCTTCCGAAGGGAAGTGTCTCCACTTGGGAGTCAATTA 2280
Db 2221 CACCCGAGCCAGCGGAGCCCTTCCGAAGGGAAGTGTCTCCACTTGGGAGTCAATTA 2280
Qy 2281 AAAGATTAGTCACTCAAGAAAGAAATCAAGTCAAAATCGGAAGAGAAAGCCGAGAGCT 2340
Db 2281 AAAGATTAGTCACTCAAGAAAGAAATCAAGTCAAAATCGGAAGAGAAAGCCGAGAGCT 2340
Qy 2341 CTAGTCTAGAGCAGTGTCTCACTGAGATCGAACCGAGTGAAGAAATCTTGGGTTTCCA 2400
Db 2341 CTAGTCTAGAGCAGTGTCTCACTGAGATCGAACCGAGTGAAGAAATCTTGGGTTTCCA 2400
Qy 2401 TTAAGAAATTCATCCCGGAGCGGGAAGAAAGGCGAGACGGGAAGCAAGACCA 2460
Db 2401 TTAAGAAATTCATCCCGGAGCGGGAAGAAAGGCGAGACGGGAAGCAAGACCA 2460
Qy 2461 CTGTGAAGACTCAGGGCAGTGGAGATAAATGAGAGCAGCCCTAATGTCCAGACCGTCG 2520
Db 2461 CTGTGAAGACTCAGGGCAGTGGAGATAAATGAGAGCAGCCCTAATGTCCAGACCGTCG 2520
Qy 2521 TGCCTCTGTCTGAGTATAATGAGTGGAGGAGAGAGATGGAAGCCAGGGGAATACGG 2580
Db 2521 TGCCTCTGTCTGAGTATAATGAGTGGAGGAGAGAGATGGAAGCCAGGGGAATACGG 2580
Qy 2581 AGCTGCCCGAGCTGTGGGGCTGTGTAGCTGTCCGAGGAGCTCAGTAAGACTCTGGTCC 2640
Db 2581 AGCTGCCCGAGCTGTGGGGCTGTGTAGCTGTCCGAGGAGCTCAGTAAGACTCTGGTCC 2640
Qy 2641 ACAGTGTAGTGTGCGAGTCAATTTGAAGGAGCAGGGCAGTCCAGTGTGGAAGCGGT 2700
Db 2641 ACAGTGTAGTGTGCGAGTCAATTTGAAGGAGCAGGGCAGTCCAGTGTGGAAGCGGT 2700
Qy 2701 CTCCTTCGTTGGATATCCGCTTCCGTAACAGAACCTTGAACACACAGCGGGAGAGCCA 2760
Db 2701 CTCCTTCGTTGGATATCCGCTTCCGTAACAGAACCTTGAACACACAGCGGGAGAGCCA 2760

QY 2761 TGCCACTGTTGAAGAGTCACTGAAAAACATCATTTGCAAGAAACTCTGTGCTCA 2820
DB 2761 TGCCACTGTTGAAGAGTCACTGAAAAACATCATTTGCAAGAAACTCTGTGCTCA 2820
QY 2821 CCCGAGCTTAAAGAGTAAAGTCCCATGACATGTGTCACTGATCCAGTAAAGATTT 2880
DB 2821 CCCGAGCTTAAAGAGTAAAGTCCCATGACATGTGTCACTGATCCAGTAAAGATTT 2880
QY 2881 TCACCTCAGAAAGCTGTGACAGCCACAGAGACCTCAAGAGCTCTCCGTACTGAAGATTA 2940
DB 2881 TCACCTCAGAAAGCTGTGACAGCCACAGAGACCTCAAGAGCTCTCCGTACTGAAGATTA 2940
QY 2941 CCGAAGCATCGGGGGCCGAAGAGACCAAGACATGTGTCCGCACTTTCCAGCTGACCTG 3000
DB 2941 CCGAAGCATCGGGGGCCGAAGAGACCAAGACATGTGTCCGCACTTTCCAGCTGACCTG 3000
QY 3001 ACTCCCGACGACCAAGAGAGCCAGATTCAAGAGGTGAGAGTGTGTCTAG 3060
DB 3001 ACTCCCGACGACCAAGAGAGCCAGATTCAAGAGGTGAGAGTGTGTCTAG 3060
QY 3061 ATACGAAAGAGAGAGCCGACAGAGCCATCTTCAAGCCGTTGACAGCAAGGTGA 3120
DB 3061 ATACGAAAGAGAGAGCCGACAGAGCCATCTTCAAGCCGTTGACAGCAAGGTGA 3120
QY 3121 AAGAAGAGTCCAGGTGCTGCAACCCAGACTGTGACAGAAACGGGGTCAAAAGCATGG 3180
DB 3121 AAGAAGAGTCCAGGTGCTGCAACCCAGACTGTGACAGAAACGGGGTCAAAAGCATGG 3180
QY 3181 AAGAAGTGAAGAGTGAAGAGAGCTCCGAAGTGTGGCTTCGGAAGAAAGAGAGAGCG 3240
DB 3181 AAGAAGTGAAGAGTGAAGAGAGCTCCGAAGTGTGGCTTCGGAAGAAAGAGAGAGCG 3240
QY 3241 TTATGCCGAAAGAGCCCGTGCAGAGCTGAGCTGAGCATTTTCACAGGGCTCTGAGA 3300
DB 3241 TTATGCCGAAAGAGCCCGTGCAGAGCTGAGCTGAGCATTTTCACAGGGCTCTGAGA 3300
QY 3301 CTGACAGAGCTACTTCAGAGAGCTTTGAAGTTCTGAAATGCAAGCAATGTGAACATG 3360
DB 3301 CTGACAGAGCTACTTCAGAGAGCTTTGAAGTTCTGAAATGCAAGCAATGTGAACATG 3360
QY 3361 TCGCAGAGTGCAGGTATTCAGAGCTCCAGAGCTGATGGAACAGGCGCCGTGAGT 3420
DB 3361 TCGCAGAGTGCAGGTATTCAGAGCTCCAGAGCTGATGGAACAGGCGCCGTGAGT 3420
QY 3421 CATCCGAAACCTTGACAGACAGTGAAGCAAAATGGAAGCACTCTTACAGATTCAGACA 3480
DB 3421 CATCCGAAACCTTGACAGACAGTGAAGCAAAATGGAAGCACTCTTACAGATTCAGACA 3480
QY 3481 CTGCAGATGGGACAGAGCAAGATGAACCATTTGACAGCCAGAGCATGAAGCCACTGACG 3540
DB 3481 CTGCAGATGGGACAGAGCAAGATGAACCATTTGACAGCCAGAGCATGAAGCCACTGACG 3540
QY 3541 CTGTAGGAGTCAAGGTCAAGAGAAAGAGGCGGCTTCTGCTAGAAAGAGAGCCCTT 3600
DB 3541 CTGTAGGAGTCAAGGTCAAGAGAAAGAGGCGGCTTCTGCTAGAAAGAGAGCCCTT 3600
QY 3601 CGACACTACTTAATTAATTTCCAGCCAGAGAAACATGGGGAAGAACAGAGATG 3660
DB 3601 CGACACTACTTAATTAATTTCCAGCCAGAGAAACATGGGGAAGAACAGAGATG 3660
QY 3661 TTCTTGAACCTACAGACAGAAAGCTTACTGCTGACAGCCGCTGCTTCTGCAAAAGACTG 3720
DB 3661 TTCTTGAACCTACAGACAGAAAGCTTACTGCTGACAGCCGCTGCTTCTGCAAAAGACTG 3720
QY 3721 AGGTGGGTCAAGAGGTGAGTTGACTGTGTGATGGAAGAAAGTCAAAAGAAAGAGAG 3780
DB 3721 AGGTGGGTCAAGAGGTGAGTTGACTGTGTGATGGAAGAAAGTCAAAAGAAAGAGAG 3780
QY 3781 AGGTGTTGTATCACTCTGGAACCAACAGTCAAAAGGCTGCTATGTGACATTTGACAGTG 3840
DB 3781 AGGTGTTGTATCACTCTGGAACCAACAGTCAAAAGGCTGCTATGTGACATTTGACAGTG 3840

QY 3841 AAGTATGGAGTGGCCGGGTGTCAAGAAAGAGAGTACTGAAGTGCAGAGCTTTAGCC 3900
DB 3841 AAGTATGGAGTGGCCGGGTGTCAAGAAAGAGAGTACTGAAGTGCAGAGCTTTAGCC 3900
QY 3901 TGGAGAGGAGATGGAAATCTGACGTTGAAAAGAGAAAGGAGACAAAGCCAGAGC 3960
DB 3901 TGGAGAGGAGATGGAAATCTGACGTTGAAAAGAGAAAGGAGACAAAGCCAGAGC 3960
QY 3961 AAGTATGAAAGAGGTGACAGAAACAGCCGCTCTTGAGCATGAAAGAACTTACGGGA 4020
DB 3961 AAGTATGAAAGAGGTGACAGAAACAGCCGCTCTTGAGCATGAAAGAACTTACGGGA 4020
QY 4021 AGCCAGTCTGACACTTGACATGCCAGCTCAGAGAGGGGAAAGGCACTGAGAAAGCTTG 4080
DB 4021 AGCCAGTCTGACACTTGACATGCCAGCTCAGAGAGGGGAAAGGCACTGAGAAAGCTTG 4080
QY 4081 GAGGAAGCCCTTCTCTCCAGAGCCAGACCAAGCAGGTTTCATAGAGTTCAAGTTCAA 4140
DB 4081 GAGGAAGCCCTTCTCTCCAGAGCCAGACCAAGCAGGTTTCATAGAGTTCAAGTTCAA 4140
QY 4141 GCTTGAACAAACAGTCACTTCAAAACAGACAGCTGTGAAAAAGTATGAAAAAGGTTG 4200
DB 4141 GCTTGAACAAACAGTCACTTCAAAACAGACAGCTGTGAAAAAGTATGAAAAAGGTTG 4200
QY 4201 TGATTTGAGAGCAGGTGAAAGTCCAGAGTGTGAGTGCACCTTATTAACAAGTGA 4260
DB 4201 TGATTTGAGAGCAGGTGAAAGTCCAGAGTGTGAGTGCACCTTATTAACAAGTGA 4260
QY 4261 AGTCTCTGCAACGGGTGGGCACTGAGACTCTTCAGATGACAGAGACACGGTACCCCTGG 4320
DB 4261 AGTCTCTGCAACGGGTGGGCACTGAGACTCTTCAGATGACAGAGACACGGTACCCCTGG 4320
QY 4321 GGCTGAGTCTCAGGCAAGATCCATCCCAATCATATGTAATCTCTGCTCTGAAAGCACC 4380
DB 4321 GGCTGAGTCTCAGGCAAGATCCATCCCAATCATATGTAATCTCTGCTCTGAAAGCACC 4380
QY 4381 TACATCTGACCTTCAAGAGGAAATTAAGCCATCCAGAGAGCCGATCAGAGGAAAGG 4440
DB 4381 TACATCTGACCTTCAAGAGGAAATTAAGCCATCCAGAGAGCCGATCAGAGGAAAGG 4440
QY 4441 ACAAGCAGATGTGCTGTCTGATGCTGACGCAAGAGAGAGTACAGCAATCGAAAAATGCC 4500
DB 4441 ACAAGCAGATGTGCTGTCTGATGCTGACGCAAGAGAGAGTACAGCAATCGAAAAATGCC 4500
QY 4501 TCAAGGCTGAACTTGAGATCTTGGAACCTTGAGATTAAGAGCAACAATGTGCTGAACG 4560
DB 4501 TCAAGGCTGAACTTGAGATCTTGGAACCTTGAGATTAAGAGCAACAATGTGCTGAACG 4560
QY 4561 TCATTCAGACGCGTTGACAGTTCGACGTAAGAAACAGCCCGGAAACATGATGCTT 4620
DB 4561 TCATTCAGACGCGTTGACAGTTCGACGTAAGAAACAGCCCGGAAACATGATGCTT 4620
QY 4621 ATGATTCACAGACCCAGTTCTTCGATGCTGACGCTTGAACAGAGGAGCCCAACAGATGCT 4680
DB 4621 ATGATTCACAGACCCAGTTCTTCGATGCTGACGCTTGAACAGAGGAGCCCAACAGATGCT 4680
QY 4681 GGAACAAAATGAAGATGCAAGATGAACCAAGTGCAGCCGACCAAGAGAGACTTGC 4740
DB 4681 GGAACAAAATGAAGATGCAAGATGAACCAAGTGCAGCCGACCAAGAGAGACTTGC 4740
QY 4741 AAGTCTGACCGTTCTGAGAGCAGGGCTGACGCTCGGAAATGCTTCCGCGCTGACAGT 4800
DB 4741 AAGTCTGACCGTTCTGAGAGCAGGGCTGACGCTCGGAAATGCTTCCGCGCTGACAGT 4800
QY 4801 TGAAGGCGCGGTGTCAAAAGTAAGATTAAGAACTGCTCTCAACCCAAAGATTCAAA 4860
DB 4801 TGAAGGCGCGGTGTCAAAAGTAAGATTAAGAACTGCTCTCAACCCAAAGATTCAAA 4860
QY 4861 AAGAGCATGTGCTGATGAGCCCTCAGCTCCAAAGCTTAAGCCAGGAGAGGCAAGTGCCT 4920
DB 4861 AAGAGCATGTGCTGATGAGCCCTCAGCTCCAAAGCTTAAGCCAGGAGAGGCAAGTGCCT 4920
QY 4921 CTGGAACCTTAACCAAGATTCCTCCAGACCAACGACCAAGCTAAACGAGAGGGCG 4980

Db 181 AGCTCCACAGAGATGCGCAGCTGTCTTCTGTACGCGCTACCTGAACAGAGATG 240
Qy 241 TCCATGTCACAGAGAGAAACAGAGAGGCGCAGAGAGAAAGTCTGTAATGAGATGTTG 300
Db 241 TCCATGTCACAGAGAGAAACAGAGAGGCGCAGAGAGAAAGTCTGTAATGAGATGTTG 300
Qy 301 GACAGCGAGAGTCAGAGAGATGTCAGAGAGAAAGACCGAGTTGAAGAAATGGCGCAACT 360
Db 301 GACAGCGAGAGTCAGAGAGATGTCAGAGAGAAAGACCGAGTTGAAGAAATGGCGCAACT 360
Qy 361 CCACAGCTGTTGAAGATATCAAGAAAGATGCGCAGAGAGACATCAGAAATTAATTGAAC 420
Db 361 CCACAGCTGTTGAAGATATCAAGAAAGATGCGCAGAGAGACATCAGAAATTAATTGAAC 420
Qy 421 AGATCCCTGCTTCAGAGAAACATATGTAAGAAATGTTACAGCTGTGATCCAGGCTA 480
Db 421 AGATCCCTGCTTCAGAGAAACATATGTAAGAAATGTTACAGCTGTGATCCAGGCTA 480
Qy 481 ATGATGTTGGCTTCAGAGAAAGTATTTAAATTTGTTGTTTAAATTCACGTTGAAGAG 540
Db 481 ATGATGTTGGCTTCAGAGAAAGTATTTAAATTTGTTGTTTAAATTCACGTTGAAGAG 540
Qy 541 ATAAATAATGAAGATCAGATCTGTCACATCTCACTGTCAGAGAGATGAAGGCGAAG 600
Db 541 ATAAATAATGAAGATCAGATCTGTCACATCTCACTGTCAGAGAGATGAAGGCGAAG 600
Qy 601 GCGCAGAAAGCTCTGTGCGAGCTGAGACCAACAGAGAGCCAGTGTGAGACTGCGCTG 660
Db 601 GCGCAGAAAGCTCTGTGCGAGCTGAGACCAACAGAGAGCCAGTGTGAGACTGCGCTG 660
Qy 661 GAGAGTCAGCATCCAGAGAAAGTGAAGCATTCACAGAGAGCAAGAGGCAAGGCA 720
Db 661 GAGAGTCAGCATCCAGAGAAAGTGAAGCATTCACAGAGAGCAAGAGGCAAGGCA 720
Qy 721 TGAACCAAGAACAGAGCGAGCAGAGAAATCCCGCTTCAAGCCGATCTGATCAAGCGCTG 780
Db 721 TGAACCAAGAACAGAGCGAGCAGAGAAATCCCGCTTCAAGCCGATCTGATCAAGCGCTG 780
Qy 781 AGAGAGAAAGCCAAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Db 781 AGAGAGAAAGCCAAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Qy 841 AATCCCGAG 900
Db 841 AATCCCGAG 900
Qy 901 AGCGTTGGGCGCGCTGCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
Db 901 AGCGTTGGGCGCGCTGCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
Qy 961 AAACTGCCAG 1020
Db 961 AAACTGCCAG 1020
Qy 1021 AGACAG 1080
Db 1021 AGACAG 1080
Qy 1081 CAGCAG 1140
Db 1081 CAGCAG 1140
Qy 1141 CCGTCAG 1200
Db 1141 CCGTCAG 1200
Qy 1201 ACCAAG 1260
Db 1201 ACCAAG 1260
Qy 1261 GAG 1320
Db 1261 GAG 1320

Db 1261 GAG 1320
Qy 1321 CTGAG 1380
Db 1321 CTGAG 1380
Qy 1381 TGAAG 1440
Db 1381 TGAAG 1440
Qy 1441 CTGAAG 1500
Db 1441 CTGAAG 1500
Qy 1501 CCTCTCAG 1560
Db 1501 CCTCTCAG 1560
Qy 1561 GCTTAAG 1620
Db 1561 GCTTAAG 1620
Qy 1621 AGCTGAG 1680
Db 1621 AGCTGAG 1680
Qy 1681 GAGAGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740
Db 1681 GAGAGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740
Qy 1741 GCGTGAAG 1800
Db 1741 GCGTGAAG 1800
Qy 1801 AG 1860
Db 1801 AG 1860
Qy 1861 TCCGAG 1920
Db 1861 TCCGAG 1920
Qy 1921 TGTCTCCAG 1980
Db 1921 TGTCTCCAG 1980
Qy 1981 AACAAAG 2040
Db 1981 AACAAAG 2040
Qy 2041 TTTGTGTGATCATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2100
Db 2041 TTTGTGTGATCATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2100
Qy 2101 GCGCAG 2160
Db 2101 GCGCAG 2160
Qy 2161 CCGGAG 2220
Db 2161 CCGGAG 2220
Qy 2221 CACCCGAG 2280
Db 2221 CACCCGAG 2280
Qy 2281 AAAAGTTAGTCACTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2340
Db 2281 AAAAGTTAGTCACTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2340
Qy 2341 CTAGTGTAG 2400
Db 2341 CTAGTGTAG 2400

Qy 2401 TTAAGAAATTCATCCCGGACGGCGGAAGAAAAGGGGCGAGACGGGAAGCAAGGCCA 2460
Db 2401 TTAAGAAATTCATCCCGGACGGCGGAAGAAAAGGGGCGAGATGGGAAGCAAGGCCA 2460
Qy 2461 CTGTGGAAAGACTCAGGGCCAGTGGAGATAAATGAGGACGACCCCTAATGTCTCCACCGCTCG 2520
Db 2461 CTGTGGAAAGACTCAGGGCCAGTGGAGATAAATGAGGACGACCCCTGATGTCCAGCCGCTCG 2520
Qy 2521 TGCCTCTGTCTGAGTATAATGACGTGGAGGGGAGAGATGGAAGCCCGAGGGGAATACGG 2580
Db 2521 TGCCTCTGTCTGAGTATGATGATGACGTGGAGGGGAGAGATGGAAGCCCGAGGGGAATGCGG 2580
Qy 2581 AGCTGCCCCAGCTCTCGGGGCTGTGTACGTGTCCGAGGAGCTCAGTAAGACTCTGTGTCC 2640
Db 2581 AGCTGCCCCAGCTCTCGGGGCTGTGTACGTGTCCGAGGAGCTCAGTAAGACTCTGTGTCC 2640
Qy 2641 ACACCTGTGAGTGTGCGAGTCAATTCATGGGACCGGGCAGTCACCAAGTGTGCGAAGCGGT 2700
Db 2641 ACACCTGTGAGTGTGCGAGTCAATTCATGGGACCGGGCAGTCACCAAGTGTGCGAAGCGGT 2700
Qy 2701 CTCCTTCGTGGATATCCGCTTCCGTAAACAGAACTCTTTGAAACACACAGCGGGAGGCCA 2760
Db 2701 CTCCTTCGTGGATATCCGCTTCCGTAAACAGAACTCTTTGAAACACACAGCGGGAGGCCA 2760
Qy 2761 TGCCACCTGTTGAAGAGGTCTACTGAAAAGACATCTTGCAGAGAACTCCTGTGTCTCA 2820
Db 2761 TGCCACCTGTTGAAGAGGTCTACTGAAAAGACATCTTGCAGAGAACTCCTGTGTCTCA 2820
Qy 2821 CCCAGAGTTTACCAGAGGGTAAAGATGCCATGACGATGGTCAACCAAGTGAAGTGATT 2880
Db 2821 CCCAGAGTTTACCAGAGGGTAAAGATGCCATGACGATGGTCAACCAAGTGAAGTGATT 2880
Qy 2881 TCACCTCAGAAGCTGTGACAGCCACAGAGACCTCAGAGGCTCTCCGTACTGAAAGAGTTA 2940
Db 2881 TCACCTCAGAAGCTGTGACAGCCACAGAGACCTCAGAGGCTCTCCGTACTGAAAGAGTTA 2940
Qy 2941 CCGAAGCATCGGGGGCGAAGAGACCAAGACATGGTGTCCGCAAGTTTCCAGCTGACTG 3000
Db 2941 CCGAAGCATCGGGGGCGAAGAGACCAAGACATGGTGTCCGCAAGTTTCCAGCTGACTG 3000
Qy 3001 ACTCCACAGACACACAGAGGAAGCCACCCAGTTCAGGAGGTAGAGNGTGGTGTCTAG 3060
Db 3001 ACTCCACAGACACACAGAGGAAGCCACCCAGTTCAGGAGGTAGAGNGTGGTGTCTAG 3060
Qy 3061 ATACAGAAAGAGGAGCGCCAGACGCGAGGCCATCTCCAAAGCCGTTGCGACACAAGGTGA 3120
Db 3061 ATACAGAAAGAGGAGCGCCAGACGCGAGGCCATCTCCAAAGCCGTTGCGACACAAGGTGA 3120
Qy 3121 AAGAGGATGCCAGGTGCTGCAACCCAGACTGTGCGAGAGAACGGGTTCAAAAGCACTGG 3180
Db 3121 AAGAGGATGCCAGGTGCTGCAACCCAGACTGTGCGAGAGAACGGGTTCAAAAGCACTGG 3180
Qy 3181 AGRAGGTTGAGGAGGTAGAGGAGACTCCGAGTGTCTGGCTTCGGAGAAAGAGAGACG 3240
Db 3181 AGRAGGTTGAGGAGGTAGAGGAGACTCCGAGTGTCTGGCTTCGGAGAAAGAGAGACG 3240
Qy 3241 TTATGCCGAAAGGACCCGTGCGAAGCTGGAGCTGAGCATCTTGGCACAGGGCTCTGAGA 3300
Db 3241 TTATGCCGAAAGGACCCGTGCGAAGCTGGAGCTGAGCATCTTGGCACAGGGCTCTGAGA 3300
Qy 3301 CTGACAGGCTTACTCCAGAGAGCCCTTGAAGTTCCTGAAGTCAACGCGCAGATGTAGACCATG 3360
Db 3301 CTGACAGGCTTACTCCAGAGAGCCCTTGAAGTTCCTGAAGTCAACGCGCAGATGTAGACCATG 3360
Qy 3361 TCGCCAGCTGCGAGGTTATCAAGTTCGAGCAGCTGATGGAAACAGGCCGCGCTCCTGAGT 3420
Db 3361 TCGCCAGCTGCGAGGTTATCAAGTTCGAGCAGCTGATGGAAACAGGCCGCGCTCCTGAGT 3420
Qy 3421 CATCCGAAACCTTCACACAGACTCAGACAAATGGAACACTCCCTTAGCAGATTTCAGACA 3480
Db 3421 CATCCGAAACCTTCACACAGACTCAGACAAATGGAACACTCCCTTAGCAGATTTCAGACA 3480

Qy 3481 CTCAGATGGGACACAGCAAGATGAAACCAATTGCACAGCCAGGACAGTAAGACCACTGCAG 3540
Db 3481 CTCAGATGGGACACACAGCAAGATGAAACCAATTGCACAGCCAGGACAGTAAGACCACTGCAG 3540
Qy 3541 CTGTCAAGGAGTCAAGGTCACAGAGAGAGAGGCGGTCTCTGTCTCAGAAAGAGAGGCTT 3600
Db 3541 CTGTCAAGGAGTCAAGGTCACAGAGAGAGAGGCGGTCTCTGTCTCAGAAAGAGAGGCTT 3600
Qy 3601 CGACACTACCTAATAATGTTTCCAGCCCCAGGAACAATCGGGGAAGAACCCAGGAAGAGATG 3660
Db 3601 CGACACTACCTAATAATGTTTCCAGCCCCAGGAAGAAATCGGGGAAGAACCCAGGAAGAGATG 3660
Qy 3661 TTCTTGAACCTACACAGCAAGAGCTTACTGCTGAGCCGCTGCCGCTTCTGCGAAAGACTG 3720
Db 3661 TTCTTGAACCTACACAGCAAGAGCTTACTGCTGAGCCGCTGCCGCTTCTGCGAAAGACTG 3720
Qy 3721 AGTGGGTCAAGAGGTGAGGTTGACTTGGTTGGATGGAGAAAAGTCAAAAGAAAGACAGG 3780
Db 3721 AGTGGGTCAAGAGGTGAGGTTGACTTGGTTGGATGGAGAAAAGTCAAAAGAAAGACAGG 3780
Qy 3781 AGTGTGTTGTACACTCTCTGGAACCAACAGTCAAAAGGCTCTGTATGTGACATATGACAGTG 3840
Db 3781 AGTGTGTTGTACACTCTCTGGAACCAACAGTCAAAAGGCTCTGTATGTGACATATGACAGTG 3840
Qy 3841 AAGTGTGGAGTGGCGGGTGTGCGGAGAAAGAGAGTACTGAAGTGCAGAGTCTTAGCC 3900
Db 3841 AAGTGTGGAGTGGCGGGTGTGCGGAGAAAGAGAGTACTGAAGTGCAGAGTCTTAGCC 3900
Qy 3901 TGGAGGAGGAGAGATGGAACCTGACGTTGAAAGAGAAAAGGGAGACAAAGCCAGAGC 3960
Db 3901 TGGAGGAGGAGAGATGGAACCTGACGTTGAAAGAGAAAAGGGAGACAAAGCCAGAGC 3960
Qy 3961 AAGTGTGAAAGAGGTGAGCAGGAAAACAGCCGCTCTCTGAGCATGAAGAAACCTACCGGA 4020
Db 3961 AAGTGTGAAAGAGGTGAGCAGGAAAACAGCCGCTCTCTGAGCATGAAGAAACCTACCGGA 4020
Qy 4021 AGCCAGTCTGACACTCTCTCCAGCATGCCAGTCTGAGAGGGGAAAGGCACTGGGAAGCCTTG 4080
Db 4021 AGCCAGTCTGACACTCTCTCCAGCATGCCAGTCTGAGAGGGGAAAGGCACTGGGAAGCCTTG 4080
Qy 4081 GAGGAAGCCCTCTCTCCAGACCAAGACAAAGCAGGTTGCATAGAGTTCAAGTTCAAA 4140
Db 4081 GAGGAAGCCCTCTCTCTCCAGACCAAGACAAAGCAGGTTGCATAGAGTTCAAGTTCAAA 4140
Qy 4141 GCCTGGACACAAACAGTCACTCAAAACAGCAGAAAGCTGTGGAAAAGGTCATAGAAACGGTTG 4200
Db 4141 GCCTGGACACAAACAGTCACTCAAAACAGCAGAAAGCTGTGGAAAAGGTCATAGAAACGGTTG 4200
Qy 4201 TGATTTTCAGAGACAGGTGAAAGTCCAGAGTGTGTAGTGCACACTTATTATACAGCTGAGA 4260
Db 4201 TGATTTTCAGAGACAGGTGAAAGTCCAGAGTGTGTAGTGCACACTTATTATACAGCTGAGA 4260
Qy 4261 AGTCCTCTGCAAAGGTTGGCCACTGGACTCTTCAGCATGCAGAGGACACGGTACCCCTGG 4320
Db 4261 AGTCCTCTGCAAAGGTTGGCCACTGGACTCTTCAGCATGCAGAGGACACGGTACCCCTGG 4320
Qy 4321 GGCTGAGTCTCAGGACAGAAATCCATCCCAATCATAGTAACTCTCTGCTCTGAAAAGCACCC 4380
Db 4321 GGCTGAGTCTCAGGACAGAAATCCATCCCAATCATAGTAACTCTCTGCTCTGAAAAGCACCC 4380
Qy 4381 TACATCTGTACCTTACAGGAGAAATAAGCGCATCCAGAGAGAGCGATCAGAGGAAGAGG 4440
Db 4381 TACATCTGTACCTTACAGGAGAAATAAGCGCATCCAGAGAGAGCGATCAGAGGAAGAGG 4440
Qy 4441 ACAAGCCAGATGCTGGTCTGTGATGCTGACGGCAAGAGAGTACAGCAATCGAAAAGTCC 4500
Db 4441 ACAAGCCAGATGCTGGTCTGTGATGCTGACGGCAAGAGAGTACAGCAATCGAAAAGTCC 4500
Qy 4501 TCAAGGCTGAACCTGAGACTCTGAACTTTCAGAGTGAAGCAACAAGATTGCTGAAAG 4560
Db 4501 TCAAGGCTGAACCTGAGACTCTGAACTTTCAGAGTGAAGCAACAAGATTGCTGAAAG 4560
Qy 4561 TCATTTCAGACAGCGGTTGACCGTTCGACGTAACAGAAACAGCCCCCGGAAACTCATGCTT 4620

```

Db 4561 TCATTGACAGCGGTTTACAGGTTTCCAGTACAGAAACAGCCCGAACTCATGCTT 4620
Qy 4621 ATGATTCACAGACCCAGGTTCTTCATGACAGGAGGAGCCCAACAGATCT 4680
Db 4621 ATGATTCACAGACCCAGGTTCTTCATGACAGGAGGAGCCCAACAGATCT 4680
Qy 4681 GGACAAAATGAAAGATCCCAAGATGAAACACCCAGTCCGAGCCCAAGAGATCTGC 4740
Db 4681 GGACAAAATGAAAGATCCCAAGATGAAACACCCAGTCCGAGCCCAAGAGATCTGC 4740
Qy 4741 AAGTCTGACCGTTCTGAGGATGGGCTCAGCCCTCGGAAATGCTTGGCCGCTTGCAGT 4800
Db 4741 AAGTCTGACCGTTCTGAGGATGGGCTCAGCCCTCGGAAATGCTTGGCCGCTTGCAGT 4800
Qy 4801 TGAAGCCCGGATGTCAAGATGAGATGAGAAAGCTGCTCTCAACCCAAAGATCCAA 4860
Db 4801 TGAAGCCCGGATGTCAAGATGAGATGAGAAAGCTGCTCTCAACCCAAAGATCCAA 4860
Qy 4861 AGAGCATGCTGCTGATGAGGCTCAGCTCAGCTCAAGCTTACCCAGAGGAGGAGGCT 4920
Db 4861 AGAGCATGCTGCTGATGAGGCTCAGCTCAGCTCAAGCTTACCCAGAGGAGGAGGCT 4920
Qy 4921 CTGGAACCTTAACCAAAATCCCAAGACCAAGGTAACCAAGGTAACCAAGGTAAC 4980
Db 4921 CTGGAACCTTAACCAAAATCCCAAGACCAAGGTAACCAAGGTAACCAAGGTAAC 4980
Qy 4981 ATCCCCCAAGGTTGAGGTTCCAGAGAAAGATGTCTACCAAGTCAAGTCAAGGTAAC 5040
Db 4981 ATCCCCCAAGGTTGAGGTTCCAGAGAAAGATGTCTACCAAGTCAAGTCAAGGTAAC 5040
Qy 5041 AGGCCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5100
Db 5041 AGGCCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5100
Qy 5101 TTAGTGTCTCATTTGATCAATCTGTAAAGACCAAGATGTAAGAAACAAATGTAAGT 5160
Db 5101 TTAGTGTCTCATTTGATCAATCTGTAAAGACCAAGATGTAAGAAACAAATGTAAGT 5160
Qy 5161 CTGCTGTGGGAGCTTGAAGACCAAGATTTCAAGAGCCATGATCCAGAGAGAGAGAG 5220
Db 5161 CTGCTGTGGGAGCTTGAAGACCAAGATTTCAAGAGCCATGATCCAGAGAGAGAGAG 5220
Qy 5221 TCCAAATGATTTCCACC 5237
Db 5221 TCCAAATGATTTCCACC 5236

```

```

; LENGTH: 5236
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. NM_057103
US-10-191-803-279

Query Match      84.5%; Score 5204.2; DB 41; Length 5236;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 5223; Conservative 0; Mismatches 13; Indels 1; Gaps 1;

Qy 1 CGGCGCGGAGAGTGAAGAGACCACTGAGCCATGGGCGAGGAGGAGTTCACCGAGAGC 60
Db 1 CGGCGCGGAGAGTGAAGAGACCACTGAGCCATGGGCGAGGAGGAGTTCACCGAGAGC 60
Qy 61 GAGCCCGGAGAGCCCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
Db 61 GAGCCCGGAGAGCCCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
Qy 121 GGGCCGAGCTGAAGCTTCGGAGACACTGAGACCCCGCGAGAGAGAGAGAGAGAGAGAG 180
Db 121 GGGCCGAGCTGAAGCTTCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
Qy 181 AGCTCCCAAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
Db 181 AGCTCCCAAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
Qy 241 TCCATGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
Db 241 TCCATGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
Qy 301 GACAGCGAGAGTCAAGAGATGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
Db 301 GACAGCGAGAGTCAAGAGATGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
Qy 361 CCAGAGCTGTGAAGATATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Db 361 CCAGAGCTGTGAAGATATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Qy 421 AGATCCCTGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
Db 421 AGATCCCTGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
Qy 481 ATGATGTGGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Db 481 ATGATGTGGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Qy 541 ATAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Db 541 ATAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Qy 601 GGGAGAGAGCTTGTGGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Db 601 GGGAGAGAGCTTGTGTGGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Qy 661 GAGAGTCAAGATCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Db 661 GAGAGTCAAGATCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Qy 721 TGAAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
Db 721 TGAAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
Qy 781 AGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Db 781 AGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Qy 841 AATCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
Db 841 AATCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
Qy 901 ACGTGGGCGGCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960

```

```

RESULT 5
US-10-191-803-279
; Sequence 279, Application US/10191803
; GENERAL INFORMATION:
; APPLICANT: MENDRICK, Donna
; APPLICANT: PORTER, Mark
; APPLICANT: JOHNSON, Kory
; APPLICANT: HIGGS, Brandon
; APPLICANT: CASTLE, Arthur
; APPLICANT: ELASHOFF, Michael
; TITLE OF INVENTION: Cardiotoxin Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5098US
; CURRENT FILING DATE: US/10/191,803
; PRIOR FILING DATE: 2002-07-10
; PRIOR FILING DATE: 2001-07-10
; PRIOR FILING DATE: 2001-07-10
; PRIOR FILING DATE: 2001-07-17
; PRIOR FILING DATE: 2001-07-17
; PRIOR FILING DATE: 2002-04-03
; PRIOR FILING DATE: 2002-05-06
; NUMBER OF SEQ ID NOS: 1140
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 279

```


Db 901 ACGTTGGCCGGCTGGCGCAAGAACACCGTCTTCAAGAAATCAAAAGAGGATGATCTGG 960
Qy 961 AAATGCCGAGAGAGAAAGAGCAAGAGGCGAGAAAAGTAGACGAGAAAGAAAGAAA 1020
Db 961 AAATGCCGAGAGAGAAAGAGCAAGAGGCGAGAAAAGTAGACGAGAAAGAAAGAAA 1020
Qy 1021 AGACAGAGCCAGCTCCGAGGAGCAGGAGCCGCGAGAGACACAGACACAGCCAGGTTGT 1080
Db 1021 AGACAGAGCCAGCTCCGAGGAGCAGGAGCCGCGAGAGACACAGACACAGCCAGGTTGT 1080
Qy 1081 CAGCAGACTACGAGAGGTGAGCTGCTCTTTGGAAAGACACAGTTGGTGGTGGTGGTGGT 1140
Db 1081 CAGCAGACTACGAGAGGTGAGCTGCTCTTTGGAAAGACACAGTTGGTGGTGGTGGTGGT 1140
Qy 1141 CGTCAGAGGAGAGTGTCTCTCTTTGGCAACGGAAGTGTGGTGGTGGTGGTGGTGGT 1200
Db 1141 CGTCAGAGGAGAGTGTCTCTCTTTGGCAACGGAAGTGTGGTGGTGGTGGTGGTGGT 1200
Qy 1201 ACCAAGAGTGTGGCAGAGGTCCACGTTGACGACCGTGGAGAACACAGAGGAGGAGG 1260
Db 1201 ACCAAGAGTGTGGCAGAGGTCCACGTTGACGACCGTGGAGAACACAGAGGAGGAGG 1260
Qy 1261 GAGGAGGAGGAGAGGCTGAAGGGGGCGTGGTGGTAGAAGGAACAGGAGAACTCTTGGCCC 1320
Db 1261 GAGGAGGAGGAGAGGCTGAAGGGGGCGTGGTGGTAGAAGGAACAGGAGAACTCTTGGCCC 1320
Qy 1321 CTGAGAAACTCTGGCTGAGCCCCCAGGAGGTCCCCAGGAAAGTGGCTGCTGAGAGCTGA 1380
Db 1321 CTGAGAAACTCTGGCTGAGCCCCCAGGAGGTCCCCAGGAAAGTGGCTGCTGAGAGCTGA 1380
Qy 1381 TGAAGACAGAGAGATGTGTCTCTGGAGGAGACCACTCAACTGACAGACCTAAGTC 1440
Db 1381 TGAAGACAGAGAGATGTGTCTCTGGAGGAGACCACTCAACTGACAGACCTAAGTC 1440
Qy 1441 CTGAAGAGAAAGCTGCCCCAAAACCCAGAAAGCATTTGTCAAGTGAGGTGAGATGCTGT 1500
Db 1441 CTGAAGAGAAAGCTGCCCCAAAACCCAGAAAGCATTTGTCAAGTGAGGTGAGATGCTGT 1500
Qy 1501 CCTCTCAGGAAAGAACTCAAGTACAGGAACTCTCTTGAAGAACTCTTCAAGTACGCTCAG 1560
Db 1501 CCTCTCAGGAAAGAACTCAAGTACAGGAACTCTCTTGAAGAACTCTTCAAGTACGCTCAG 1560
Qy 1561 GCTTAAAGAACTCTCTGGGAAAGACAGAAAGGGGAAACGAGGAGGTGGGGAGACGAA 1620
Db 1561 GCTTAAAGAACTCTCTGGGAAAGACAGAAAGGGGAAACGAGGAGGTGGGGAGACGAA 1620
Qy 1621 AGCTCGAGAAATCAACACATTCACACCGAAATCCACAGAGAGTCTGATCAGCAGAGG 1680
Db 1621 AGCTCGAGAAATCAACACATTCACACCGAAATCCACAGAGAGTCTGATCAGCAGAGG 1680
Qy 1681 GAGAGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740
Db 1681 GAGAGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740
Qy 1741 CGCTGGAAGCAACCCAGGATGGGAGCTGAGGAGGAACTACTTCCGATGGAGAGAA 1800
Db 1741 CGCTGGAAGCAACCCAGGATGGGAGCTGAGGAGGAACTACTTCCGATGGAGAGAA 1800
Qy 1801 AGACAGAGGAGTCACTCTCTGGGATCTCTTCAAAAGATGCTGACACCCCAAGAAACGG 1860
Db 1801 AGACAGAGGAGTCACTCTCTGGGATCTCTTCAAAAGATGCTGACACCCCAAGAAACGG 1860
Qy 1861 TCCGAAGACCTTCTGAGAGTGAACGAGGAGAGCTGGAGAGGTCAAGAGCGCCACT 1920
Db 1861 TCCGAAGACCTTCTGAGAGTGAACGAGGAGAGCTGGAGAGGTCAAGAGCGCCACT 1920
Qy 1921 TGTCTTCACTGATAGCAGTGTGAGAAATGCAAGATGAAAGTCAAACTTTTGGTGGG 1980
Db 1921 TGTCTTCACTGATAGCAGTGTGAGAAATGCAAGATGAAAGTCAAACTTTTGGTGGG 1980
Qy 1981 AACAAAAGCCAGAGGAAACCAAGCGTAGGGTGGATCTTCAAGTCTTGGGAAGCACTGA 2040
Db 1981 AACAAAAGCCAGAGGAAACCAAGCGTAGGGTGGATCTTCAAGTCTTGGGAAGCACTGA 2040

Qy 2041 TTTGTGTCCGATCATCAAGAAAGAGAGCAAGAGGAGCATCTCTTTCAGATGATGAAGGAG 2100
Db 2041 TTTGTGTCCGATCATCAAGAAAGAGAGCAAGAGGAGCATCTCTTTCAGATGATGAAGGAG 2100
Qy 2101 GGCACAGGACACTGGGAGGGACAGTCCACAGACAGAGGAGGCGCAGCAAGACAAAG 2160
Db 2101 GGCACAGGACACTGGGAGGGACAGTCCACAGACAGAGGAGGCGCAGCAAGACAAAG 2160
Qy 2161 CCGGAAACAGACGCTGTTCTGCGCAGCACCCAGGAGCAGGACCAAGCGCAAGAGTTCTCT 2220
Db 2161 CCGGAAACAGACGCTGTTCTGCGCAGCACCCAGGAGCAGGACCAAGCGCAAGAGTTCTCT 2220
Qy 2221 CACCCGAGCCAGCGGGAGGCGCTTCCGAGGGGAAGGTGTCTCCACTTGGGAGTCAATTA 2280
Db 2221 CACCCGAGCCAGCGGGAGGCGCTTCCGAGGGGAAGGTGTCTCCACTTGGGAGTCAATTA 2280
Qy 2281 AAAGATTAGTCACTCCAAAGAAAATAATCCAAAGTCAAAATCTGGAAGAGAAAGCCGAGACT 2340
Db 2281 AAAGATTAGTCACTCCAAAGAAAATAATCCAAAGTCAAAATCTGGAAGAGAAAGCCGAGACT 2340
Qy 2341 CTAGTGTAGAGCAGTGTCTCCACTGAGATCGAAACCGAGTAGAGAGAAATCTTTGGTTTCCA 2400
Db 2341 CTAGTGTAGAGCAGTGTCTCCACTGAGATCGAAACCGAGTAGAGAGAAATCTTTGGTTTCCA 2400
Qy 2401 TTAAGAAATTCATCCCGGACCGCGGAGAAAGAGGCGAGACGGGAGCAAGAAACAAAGCCA 2460
Db 2401 TTAAGAAATTCATCCCGGACCGCGGAGAAAGAGGCGAGATGGGAGCAAGAAACAAAGCCA 2460
Qy 2461 CTGTGGAAGACTCAGGCGCAGTGGAGATAATGAGGACGACCTTAATGTCGCCAGCTCG 2520
Db 2461 CTGTGGAAGACTCAGGCGCAGTGGAGATAATGAGGACGACCTTAATGTCGCCAGCTCG 2520
Qy 2521 TGCTCTGTCTGAGTATAATGTCAGTGGAGAGGAGAGATGGAAGCCAGGGGAAATACGG 2580
Db 2521 TGCTCTGTCTGAGTATAATGTCAGTGGAGAGGAGAGATGGAAGCCAGGGGAAATACGG 2580
Qy 2581 AGTGCCTCCAGCTGCTGGGGCTGTGTAACGAGTCCGAGGAGCTCAGTAAGACTCTGCTCC 2640
Db 2581 AGTGCCTCCAGCTGCTGGGGCTGTGTAACGAGTCCGAGGAGCTCAGTAAGACTCTGCTCC 2640
Qy 2641 ACCTGTGAGTGTGCGACTCATTTGATGGGACCCAGGGCAGTCAACAGTGTCCGAGAGCGGT 2700
Db 2641 ACCTGTGAGTGTGCGACTCATTTGATGGGACCCAGGGCAGTCAACAGTGTCCGAGAGCGGT 2700
Qy 2701 CTCTCTCGTGATATCCGCTTCCGTAACAGAACTCTTTGAACACACAGCGGGAGAGCCA 2760
Db 2701 CTCTCTCGTGATATCCGCTTCCGTAACAGAACTCTTTGAACACACAGCGGGAGAGCCA 2760
Qy 2761 TGCCACCTGTTGAAGAGGTCACTGAAAGAGACATCAATTGAGAGAAACTCTCTGCTCA 2820
Db 2761 TGCCACCTGTTGAAGAGGTCACTGAAAGAGACATCAATTGAGAGAAACTCTCTGCTCA 2820
Qy 2821 CCAGAGCTTACAGAGGTTAAAGATGCCATGACGACATGGTCCAGTCAAGTGAAGTGAAT 2880
Db 2821 CCAGAGCTTACAGAGGTTAAAGATGCCATGACGACATGGTCCAGTCAAGTGAAGTGAAT 2880
Qy 2881 TCACCTCAGAGTGTGACAGCCACAGAGACCTCAGAGGCTCTCCGTAACAGAGAGTGA 2940
Db 2881 TCACCTCAGAGTGTGACAGCCACAGAGACCTCAGAGGCTCTCCGTAACAGAGAGTGA 2940
Qy 2941 CCGAAGCATCCGGGGCGGAGAGACCAACAGATGGTGTCCGAGTGTTCAGCTGACTG 3000
Db 2941 CCGAAGCATCCGGGGCGGAGAGACCAACAGATGGTGTTCAGCTGACTG 3000
Qy 3001 ACTCCCGAGACACACAGAGAGAGCCAGTTCAGAGGTTAGAGAGTGGTGGTAG 3060
Db 3001 ACTCCCGAGACACACAGAGAGAGCCAGTTCAGAGGTTAGAGAGTGGTGGTAG 3060
Qy 3061 ATACAGAAAGAGGAGCGCCAGAGCGAGGCGATCTCCAAAGCGTTGCGAGCAAGGTGA 3120
Db 3061 ATACAGAAAGAGGAGCGCCAGAGCGAGGCGATCTCCAAAGCGTTGCGAGCAAGGTGA 3120

QY 3121 AAGAGAGTCCAGGTGCTGCAACCCAGACTGTGACAGAGAACGGGGTCAAAAGCACTGG 3180
| | | | |
Db 3121 AAGAGAGTCCAGGTGCTGCAACCCAGACTGTGACAGAGAACGGGGTCAAAAGCACTGG 3180
| | | | |
QY 3181 AGAAGTTGAGAGGTAGAGAGAGACTCCGAAGTGTGCTTCCGAGAGAAAGAGAGAGCG 3240
| | | | |
Db 3181 AGAAGTTGAGAGGTAGAGAGAGACTCCGAAGTGTGCTTCCGAGAGAAAGAGAGAGCG 3240
| | | | |
QY 3241 TTATGCCGAAGAGACCCCGTGCAGAGAGCTGGAGCTGAGCATCTTGACAGAGGGCTTGAGA 3300
| | | | |
Db 3241 TTATGCCGAAGAGACCCCGTGCAGAGAGCTGGAGCTGAGCATCTTGACAGAGGGCTTGAGA 3300
| | | | |
QY 3301 CTGACAGGCTACTCCAGAGAGCCCTTGAAGTTCCTGAAGTCAAGCCAGATGAGACATG 3360
| | | | |
Db 3301 CTGACAGGCTACTCCAGAGAGCCCTTGAAGTTCCTGAAGTCAAGCCAGATGAGACATG 3360
| | | | |
QY 3361 TCGCACAAGTGTGCTGCAAGCTCCAGAGCTGATGGAACAGGCGGCTGAGT 3420
| | | | |
Db 3361 TCGCACAAGTGTGCTGCAAGCTCCAGAGCTGATGGAACAGGCGGCTGAGT 3420
| | | | |
QY 3421 CATCCGAACCTTGAACAGAGTGAACCAATTGAGAGCACTCCCTTACAGATTCAGACA 3480
| | | | |
Db 3421 CATCCGAACCTTGAACAGAGTGAACCAATTGAGAGCACTCCCTTACAGATTCAGACA 3480
| | | | |
QY 3481 CTGACAGATGGGACAGAGAGTGAACCAATTGAGAGCACTCCCTTACAGATTCAGACA 3540
| | | | |
Db 3481 CTGACAGATGGGACAGAGAGTGAACCAATTGAGAGCACTCCCTTACAGATTCAGACA 3540
| | | | |
QY 3541 CTGTCAGGCTGACAGGTTCACAGAGAGAGGCGGCTTACTGCTCAGAAAGAGAGCCCTT 3600
| | | | |
Db 3541 CTGTCAGGCTGACAGGTTCACAGAGAGAGGCGGCTTACTGCTCAGAAAGAGAGCCCTT 3600
| | | | |
QY 3601 CGACACTACTAATATATGTTCACGCCACAGAGAAACATGGGGAGAGACAGAGATG 3660
| | | | |
Db 3601 CGACACTACTAATATATGTTCACGCCACAGAGAAACATGGGGAGAGATG 3660
| | | | |
QY 3661 TTCTTGAACCTTACAGAGAGAGCTTACTGCTGAGAGCGGCTGCTGAGAGAGAGCTG 3720
| | | | |
Db 3661 TTCTTGAACCTTACAGAGAGAGCTTACTGCTGAGAGCGGCTGCTGAGAGAGAGCTG 3720
| | | | |
QY 3721 AGGTGGGTCAAGAGAGGTGAGTGTGACTGTTGATGAGAGAAATTCAGAGAGAGAG 3780
| | | | |
Db 3721 AGGTGGGTCAAGAGAGGTGAGTGTGACTGTTGATGAGAGAAATTCAGAGAGAGAG 3780
| | | | |
QY 3781 AGGTGGGTCAAGAGAGGTGAGTGTGACTGTTGATGAGAGAAATTCAGAGAGAGAG 3840
| | | | |
Db 3781 AGGTGGGTCAAGAGAGGTGAGTGTGACTGTTGATGAGAGAAATTCAGAGAGAGAG 3840
| | | | |
QY 3841 AAGTATGGAGTGGCCGGGTGTCAAGAAAGAGAGATTCGAATGACAGATCTTACGC 3900
| | | | |
Db 3841 AAGTATGGAGTGGCCGGGTGTCAAGAAAGAGAGATTCGAATGACAGATCTTACGC 3900
| | | | |
QY 3901 TGGAGAGAGAGAGATGAGAACTGACGTTGAAAGAGAGAGAGAGAGAGAGAGAGAGC 3960
| | | | |
Db 3901 TGGAGAGAGAGAGATGAGAACTGACGTTGAAAGAGAGAGAGAGAGAGAGAGAGAGC 3960
| | | | |
QY 3961 AAGTATGGAGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4020
| | | | |
Db 3961 AAGTATGGAGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4020
| | | | |
QY 4021 AGCCAGTCTGACACTTGAACATGCCAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAG 4080
| | | | |
Db 4021 AGCCAGTCTGACACTTGAACATGCCAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAG 4080
| | | | |
QY 4081 GAG 4140
| | | | |
Db 4081 GAG 4140
| | | | |
QY 4141 GCTTGAAG 4200
| | | | |
Db 4141 GCTTGAAG 4200
| | | | |
QY 4201 TGATTTCAAG 4260
| | | | |

Db 4201 TGATTTCAAG 4260
| | | | |
QY 4261 AGTCTCTGACAG 4320
| | | | |
Db 4261 AGTCTCTGACAG 4320
| | | | |
QY 4321 GGCCTGAGTCTCAG 4380
| | | | |
Db 4321 GGCCTGAGTCTCAG 4380
| | | | |
QY 4381 TACATCTGACCTTCAAG 4440
| | | | |
Db 4381 TACATCTGACCTTCAAG 4440
| | | | |
QY 4441 ACAAGCAGATGCTGTGCTTGAATCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4500
| | | | |
Db 4441 ACAAGCAGATGCTGTGCTTGAATCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4500
| | | | |
QY 4501 TCAAGGCTGAACCTGAGATCTGGAACCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4560
| | | | |
Db 4501 TCAAGGCTGAACCTGAGATCTGGAACCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4560
| | | | |
QY 4561 TCATTGAG 4620
| | | | |
Db 4561 TCATTGAG 4620
| | | | |
QY 4621 ATGATTCAG 4680
| | | | |
Db 4621 ATGATTCAG 4680
| | | | |
QY 4681 GGACAAAATGAAGATGAG 4740
| | | | |
Db 4681 GGACAAAATGAAGATGAG 4740
| | | | |
QY 4741 AAGTCTGACAG 4800
| | | | |
Db 4741 AAGTCTGACAG 4800
| | | | |
QY 4801 TGAAG 4860
| | | | |
Db 4801 TGAAG 4860
| | | | |
QY 4861 AG 4920
| | | | |
Db 4861 AG 4920
| | | | |
QY 4921 CTGGAACCTTACCAAG 4980
| | | | |
Db 4921 CTGGAACCTTACCAAG 4980
| | | | |
QY 4981 ATCCCCCAAG 5040
| | | | |
Db 4981 ATCCCCCAAG 5040
| | | | |
QY 5041 AGGCCAG 5100
| | | | |
Db 5041 AGGCCAG 5100
| | | | |
QY 5101 TTAAGTGTCTATTTGATCACTGTGTAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5160
| | | | |
Db 5101 TTAAGTGTCTATTTGATCACTGTGTAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5160
| | | | |
QY 5161 CTGCTGTGGAG 5220
| | | | |
Db 5161 CTGCTGTGGAG 5220
| | | | |
QY 5221 TCCATGATTTCCAGCC 5237
| | | | |
Db 5221 TCCATGATTTCCAGCC 5237
| | | | |

RESULT 6

US-08-978-277A-3
; Sequence 3, Application US/08978277A
; GENERAL INFORMATION:
; APPLICANT: Gelman, Irwin H.
; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10112-0228
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/978,277A
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/665,401
; FILING DATE: 18-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Richard S
; REGISTRATION NUMBER: 26,154
; REFERENCE/DOCKET NUMBER: A30558 - 165/34008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-408-2558
; TELEFAX: 212-765-2519
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5200 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
US-08-978-277A-3

Query Match 84.4%; Score 5200; DB 13; Length 5200;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	33	ATGGCGCAGGAGCTTCCACAGCAGCGAGCCCGCAGCAGCGCGGGGAGCGACACG	92
Db	1	ATGGCGCAGGAGCTTCCACAGCAGCGAGCCCGCAGCAGCGCGGGGAGCGACACG	60
Qy	93	CCGAGCAGCTGGTGTCTCAGTGGCCATGGCGCCGAGCTGGAAGCCCTCGGGAGCAGTGA	152
Db	61	CCGAGCAGCTGGTGTCTCAGTGGCCATGGCGCCGAGCTGGAAGCCCTCGGGAGCAGTGA	120
Qy	153	GACCCCGCGAGCGGACCCCGCCACCAAGCTCCACAGAAATGGCCAGCTGTCTTCT	212
Db	121	GACCCCGCGAGCGGACCCCGCCACCAAGCTCCACAGAAATGGCCAGCTGTCTTCT	180
Qy	213	GTCACCGGCTAGCTGAACAAGGAGTGTCCATGTCCAGAGGAAACCCAGGCGGCGAG	272
Db	181	GTCACCGGCTAGCTGAACAAGGAGTGTCCATGTCCAGAGGAAACCCAGGCGGCGAG	240
Qy	273	GAGGAAGAAGTCTGTGATGAGGATGTGGACAGCGAGGTCAAGAATGTGAGAGAAA	332
Db	241	GAGGAAGAAGTCTGTGATGAGGATGTGGACAGCGAGGTCAAGAATGTGAGAGAAA	300
Qy	333	GACCCAGTTGAAGAAATGGCGCCAACTCCACAGCTGTTGAAGATATCAAAAGGATGG	392
Db	301	GACCCAGTTGAAGAAATGGCGCCAACTCCACAGCTGTTGAAGATATCAAAAGGATGG	360

Qy	393	CAGGAGGAGACATCAGAAATAATTGAACAGATCCCTGCTTTCAGAAAAACAATGTGGAAGAA	452
Db	361	CAGGAGGAGACATCAGAAATAATTGAACAGATCCCTGCTTTCAGAAAAACAATGTGGAAGAA	420
Qy	453	ATGTCACAGCCTGCTGAGTCCAGGCTAATGATGTGGCTTCAAGAAAGTATTTAAATTT	512
Db	421	ATGTCACAGCCTGCTGAGTCCAGGCTAATGATGTGGCTTCAAGAAAGTATTTAAATTT	480
Qy	513	GTGTGTTTTAAATTCACGGTGAAGAGGATAAAATGAAAAGTCAGATACTGTCCAACTA	572
Db	481	GTGTGTTTTAAATTCACGGTGAAGAGGATAAAATGAAAAGTCAGATACTGTCCAACTA	540
Qy	573	CTCACTGTCAAGAAGGATGAAGGCGGAAGGCGAGAGCTCTGTTCGAGCTCGAGACCCAC	632
Db	541	CTCACTGTCAAGAAGGATGAAGGCGGAAGGCGAGAGCTCTGTTCGAGCTCGAGACCCAC	600
Qy	633	CAGGAGCCCGAGTGTGAGACTCCCGTCGGAGAGTCAGCATCCAAAGAAAAGTGAAGTGAAG	692
Db	601	CAGGAGCCCGAGTGTGAGACTCCCGTCGGAGAGTCAGCATCCAAAGAAAAGTGAAGTGAAG	660
Qy	693	CAATCCACAGAGAACAAGAGGACCCCTGAAGCAAGAACAGAGCAGCAGACAGAAATCCCCC	752
Db	661	CAATCCACAGAGAACAAGAGGACCCCTGAAGCAAGAACAGAGCAGCAGACAGAAATCCCCC	720
Qy	753	CTTCAAGCCGGAATCTGATCAAGCGGCTGAGGAAGAGCCAAAGATGAAGGAGAGAAAAA	812
Db	721	CTTCAAGCCGGAATCTGATCAAGCGGCTGAGGAAGAGCCAAAGATGAAGGAGAGAAAAA	780
Qy	813	CAAGAGAAGAGCCCAACCAAGTCCCAAGATCCCGAGCAGCCCAAGTCAACAGTGAAGACA	872
Db	781	CAAGAGAAGAGCCCAACCAAGTCCCAAGATCCCGAGCAGCCCAAGTCAACAGTGAAGACA	840
Qy	873	ACATCTTCTTCAAGAAAGTCTTCACTCAGCGTGGCGCGGCTGGCGCAAGAGACACG	932
Db	841	ACATCTTCTTCAAGAAAGTCTTCACTCAGCGTGGCGCGGCTGGCGCAAGAGACACG	900
Qy	933	TTCAAGAAATCAAAAGAGGATGATCTGGAATCTGCCAGAAAGAGAAAGGAGCAAGAGGCA	992
Db	901	TTCAAGAAATCAAAAGAGGATGATCTGGAATCTGCCAGAAAGAGAAAGGAGCAAGAGGCA	960
Qy	993	GAAGAGTAGACGAGGAAGAAAGAAAGACAGAGCCAGCCTCGGAGGAGCAGGAGCGG	1052
Db	961	GAAGAGTAGACGAGGAAGAAAGAAAGACAGAGCCAGCCTCGGAGGAGCAGGAGCGG	1020
Qy	1053	GCAGAGACACAGCAGCCAGGTTGTCTCAGCAGACTACGAGAGGTGAGCTGCTTTG	1112
Db	1021	GCAGAGACACAGCAGCCAGGTTGTCTCAGCAGACTACGAGAGGTGAGCTGCTTTG	1080
Qy	1113	GAAGACACAGGTTGGTGACCTCGAGGCGATCGTCAGAGGAGAAAGTGCTCTCTTTGGCAACG	1172
Db	1081	GAAGACACAGGTTGGTGACCTCGAGGCGATCGTCAGAGGAGAAAGTGCTCTCTTTGGCAACG	1140
Qy	1173	GAAGTGTTTGATGAGAAGATGGAAGCCCAACCAAGAGTTGTTTCAGAGGTCCACGTGAGC	1232
Db	1141	GAAGTGTTTGATGAGAAGATGGAAGCCCAACCAAGAGTTGTTTCAGAGGTCCACGTGAGC	1200
Qy	1233	ACCGTGGAGAACACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1292
Db	1201	ACCGTGGAGAACACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1260
Qy	1293	GTAGAAGGAACAGGAGAAATCTTTGCCCTCGTGAAGAACTCGCTGAGCCCGGAGGAGTCC	1352
Db	1261	GTAGAAGGAACAGGAGAAATCTTTGCCCTCGTGAAGAACTCGCTGAGCCCGGAGGAGTCC	1320
Qy	1353	CAGGAAGCTGAGCCTGCTGAGGAGCTGATGAAGAGCAGAGAGATGTGTCTCTGAGGGA	1412
Db	1321	CAGGAAGCTGAGCCTGCTGAGGAGCTGATGAAGAGCAGAGAGATGTGTCTCTGAGGGA	1380
Qy	1413	GACCACTCAACTGACAGACCTTAAGTCTTGAAGAGAGACGCTGCCCAAAACACCCAGAA	1472
Db	1381	GACCACTCAACTGACAGACCTTAAGTCTTGAAGAGAGACGCTGCCCAAAACACCCAGAA	1440

Db 3601 GAACATGGGAGAACCCAGGAGAGATGTTCTTGAACCTACACAGCAAGACCTTACTGCT 3660
QY 3693 GCAGCCGTGCCGTTCTGGCAAGACTGAGGTGGGTCAAGAGGGTGAGTTGACTGGTTG 3752
Db 3661 GCAGCCGTGCCGTTCTGGCAAGACTGAGGTGGGTCAAGAGGGTGAGTTGACTGGTTG 3720
QY 3753 GATGGAGAAAAGTCAAGAGACAGAGGTGTTTGTACACTCTGACCCCAACAGTCAA 3812
Db 3721 GATGGAGAAAAGTCAAGAGAACAGAGGTGTTTGTACACTCTGACCCCAACAGTCAA 3780
QY 3813 AAGCTCTCTGATGTGACATATGACAGTGAAGTGAATGGAGTGGCCGGGTGTCAGGAAAAG 3872
Db 3781 AAGCTCTCTGATGTGACATATGACAGTGAAGTGAATGGAGTGGCCGGGTGTCAGGAAAAG 3840
QY 3873 GAGAGTACTGAAGTGCAGAGTCTTAGCCTTGAGGAGGAGAGATGGAAAATGACGTTGAA 3932
Db 3841 GAGAGTACTGAAGTGCAGAGTCTTAGCCTTGAGGAGGAGAGATGGAAAATGACGTTGAA 3900
QY 3933 AAGGAGAAAAGGAGACAAAGCCAGACAGTGAAGTGAAGAGGTGAGCAGGAAACAGCC 3992
Db 3901 AAGGAGAAAAGGAGACAAAGCCAGACAGTGAAGTGAAGAGGTGAGCAGGAAACAGCC 3960
QY 3993 GCTCCTGAGCATGAAGAACTTACGGGAAGCAGTCTCTGACACTTTGACATGCCCAGCTCA 4052
Db 3961 GCTCCTGAGCATGAAGAACTTACGGGAAGCAGTCTCTGACACTTTGACATGCCCAGCTCA 4020
QY 4053 GAGAGGGGAAGGACACTGGGAAGCCTTGGAGGAAGCCCTTCTCTCCAGACCAAGACAAA 4112
Db 4021 GAGAGGGGAAGGACACTGGGAAGCCTTGGAGGAAGCCCTTCTCTCCAGACCAAGACAAA 4080
QY 4113 GCAGTTGCATAGAGTTCAAGTTCAAAGCTGAGACACACAGTCACTCAAACAGCAGAA 4172
Db 4081 GCAGTTGCATAGAGTTCAAGTTCAAAGCTGAGACACACAGTCACTCAAACAGCAGAA 4140
QY 4173 GCTGTGAAAAGGTCTAGAAAAGTGTGATTTTCAGACAGAGTGAAGTCCAGAGTGT 4232
Db 4141 GCTGTGAAAAGGTCTAGAAAAGTGTGATTTTCAGACAGAGTGAAGTCCAGAGTGT 4200
QY 4233 GTAGGTGCACACTTATTACAGCTGAGAAGTCTCTGCAACGGGTGGCCACTGGACTCTT 4292
Db 4201 GTAGGTGCACACTTATTACAGCTGAGAAGTCTCTGCAACGGGTGGCCACTGGACTCTT 4260
QY 4293 CAGCATGCAGAGACAGCTACCCCTGGGCTGAGTCTCAGGAGAGATCCATCCCAATC 4352
Db 4261 CAGCATGCAGAGACAGCTACCCCTGGGCTGAGTCTCAGGAGAGATCCATCCCAATC 4320
QY 4353 ATAGTAACTCTCTGCTGAAAAGCACCCTACATCTTACAGACAGTGAAGTAAAGCGCA 4412
Db 4321 ATAGTAACTCTCTGCTGAAAAGCACCCTACATCTTACAGACAGTGAAGTAAAGCGCA 4380
QY 4413 TCCAGAGAGAGCGATCAGAGGAGAGGACAAAGCCAGATGCTGCTCTGATGCTGACGGC 4472
Db 4381 TCCAGAGAGAGCGATCAGAGGAGAGGACAAAGCCAGATGCTGCTCTGATGCTGACGGC 4440
QY 4473 AAGGAGGTACAGCAATCGAAAAGTCTCAAGGCTGAACCTGAGATCTTGGAACTTGAG 4532
Db 4441 AAGGAGGTACAGCAATCGAAAAGTCTCAAGGCTGAACCTGAGATCTTGGAACTTGAG 4500
QY 4533 AGTAAGAGCAACAAGATTGCTGMAAGTCAATTCAGACAGCCGTTGACCACTTCGACGT 4592
Db 4501 AGTAAGAGCAACAAGATTGCTGMAAGTCAATTCAGACAGCCGTTGACCACTTCGACGT 4560
QY 4593 ACAGAAAACAGCCCGCAAACTCATGTTATGATTTCAGACAGCCAGGTTCTCTGATGCGAGG 4652
Db 4561 ACAGAAAACAGCCCGCAAACTCATGTTATGATTTCAGACAGCCAGGTTCTCTGATGCGAGG 4620
QY 4653 CTTGACAGCAGGAGGCCCAACAGATGCTGGAACAAAATGAAGATGCAAGATGAACAC 4712
Db 4621 CTTGACAGCAGGAGGCCCAACAGATGCTGGAACAAAATGAAGATGCAAGATGAACAC 4680
QY 4713 CCAGTCCCGCAGCCCGCAGAGGAGTTCGAAGTCTGACCGCTTGGAGGATGGGCTCAG 4772
Db 4681 CCAGTCCCGCAGCCCGCAGAGGAGTTCGAAGTCTGACCGCTTGGAGGATGGGCTCAG 4740

RESULT 7

PCT-US97-06830-1

; Sequence 1, Application PC/TUS9706830

; GENERAL INFORMATION:

; APPLICANT: Gelman, Irwin H.

; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE

; NUMBER OF SEQUENCES: 20

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond

; STREET: 30 Rockefeller Plaza

; CITY: New York

; STATE: NY

; COUNTRY: USA

; ZIP: 10112-0228

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq Version 1.5

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US97/06830

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/665,401

; FILING DATE: 18-JUN-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Clark, Richard S

; REGISTRATION NUMBER: 26,154

; REFERENCE/DOCKET NUMBER: A30558 - 165/34008

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212-408-2558

; TELEFAX: 212-765-2519

; TELEX:

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 5134 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

Db 1968 TTTCCAGCTGACTGCCAGACACCACAGAGGAGCCACCCAGTTTCAGAGGTAG 2027
Qy 3046 AGAGTGTGTCTAGATACAGAGAAGAGAGGCCACAGAGGCCATCTCTCAAGCCG 3105
Db 2028 AGGTGTGTGTCTAGATACAGAGAAGAGAGGCCACAGAGGCCATCTCTCAAGCCG 2087
Qy 3106 TTGCAGACAAGGTGAAGAGAGTCCAGAGTGCCTGCAACCCAGACGTGTCAGAGAACGG 3165
Db 2088 TTGCAGACAAGGTGAAGAGAGTCCAGAGTGCCTGCAACCCAGACGTGTCAGAGAACGG 2147
Qy 3166 GGTCAAAAGCACTGGAGAAGTTGAGAGGTGAGAGGAGTCCGAAAGTGTGGCTTCGG 3225
Db 2148 GGTCAAAAGCACTGGAGAAGTTGAGAGGTGAGAGGAGTCCGAAAGTGTGGCTTCGG 2207
Qy 3226 AGAAGAGAGAGGAGTTATGCCGAAAGGACCCGTGCAAGAGCTGGAGCTGAGCATCTTG 3285
Db 2208 AGAAGAGAGAGGAGTTATGCCGAAAGGACCCGTGCAAGAGCTGGAGCTGAGCATCTTG 2267
Qy 3286 CACAGGCTCTGAGACTGGACAGGCTACTCCAGAGAGCTTGAAGTTCCCTGAGTCAAG 3345
Db 2268 CACAGGCTCTGAGACTGGACAGGCTACTCCAGAGAGCTTGAAGTTCCCTGAGTCAAG 2327
Qy 3346 CAGATGTAGACCACTGCTGCCACAGTCCAGAGTTATCAAGCTCCAGCAGCTGATGGAACAGG 3405
Db 2328 CAGATGTAGACCACTGCTGCCACAGTCCAGAGTTATCAAGCTCCAGCAGCTGATGGAACAGG 2387
Qy 3406 CCGTGGCCCTGAGTCAATCCGAAACCTTTGACAGACAGTGAACAATGGAAGCACTCCCT 3465
Db 2388 CCGTGGCCCTGAGTCAATCCGAAACCTTTGACAGACAGTGAACAATGGAAGCACTCCCT 2447
Qy 3466 TAGCAGATTCAGACACTGCAGATGGGACACAGCAGATGAACATTTGACAGCCAGACA 3525
Db 2448 TAGCAGATTCAGACACTGCAGATGGGACACAGCAGATGAACATTTGACAGCCAGACA 2507
Qy 3526 GTAAAGCCACTGCAGCTGTCAAGCAGTCAAGGTTCAGAGAGAGAGCGCTACTGCTC 3585
Db 2508 GTAAAGCCACTGCAGCTGTCAAGCAGTCAAGGTTCAGAGAGAGAGCGCTACTGCTC 2567
Qy 3586 AGAAGAGGAGCTTCGACACTACTTAATATGTTTCAGGCCACAGGAAGACATGGGGAAG 3645
Db 2568 AGAAGAGGAGCTTCGACACTACTTAATATGTTTCAGGCCACAGGAAGACATGGGGAAG 2627
Qy 3646 AACAGAGAGATGTTCTTGAACCTACAGCAGAGAGCTTACTGCTGCAGCCGTCGCCG 3705
Db 2628 AACAGAGAGATGTTCTTGAACCTACAGCAGAGAGCTTACTGCTGCAGCCGTCGCCG 2687
Qy 3706 TTCTGGCAAGACTGAGTGGGTCAAGAGGTGAGTTGACTGTTGGATGGAGAAAG 3765
Db 2688 TTCTGGCAAGACTGAGTGGGTCAAGAGGTGAGTTGACTGTTGGATGGAGAAAG 2747
Qy 3766 TCAAGAGAACAGGAGGTGTTGTACACTCTGGACCCACAGTCAAAAGGCTGCTGATG 3825
Db 2748 TCAAGAGAACAGGAGGTGTTGTACACTCTGGACCCACAGTCAAAAGGCTGCTGATG 2807
Qy 3826 TGACATATGACATGAGTGAAGTGGAGTGGCCGGGTGTCAGGAAAGAGAGTACTGAAG 3885
Db 2808 TGACATATGACATGAGTGAAGTGGAGTGGCCGGGTGTCAGGAAAGAGAGTACTGAAG 2867
Qy 3886 TGACAGTCTTAGCTCGAGAGGAGAGATGGAAACCTGACGTTGAAAGAGAGAAAGGG 3945
Db 2868 TGACAGTCTTAGCTCGAGAGGAGAGATGGAAACCTGACGTTGAAAGAGAGAAAGGG 2927
Qy 3946 AGACAAAGCCAGAGCAAGTGAAGAGGTGAGCAGGAGAAACAGCCGCTCCTGAGCATG 4005
Db 2928 AGACAAAGCCAGAGCAAGTGAAGAGGTGAGCAGGAGAAACAGCCGCTCCTGAGCATG 2987
Qy 4006 AAGGAACCTACGGAAGCCAGTCTGACATTTGACATGCCAGCTCAGAGAGGGGAAGG 4065
Db 2988 AAGGAACCTACGGAAGCCAGTCTGACATTTGACATGCCAGCTCAGAGAGGGGAAGG 3047
Qy 4066 CACTGGAGAGCTTGGAGGAGCCCTTCTCTCCAGACCAAGACAAAGAGGTGTCATAG 4125
Db 3048 CACTGGAGAGCTTGGAGGAGCCCTTCTCTCTCCAGACCAAGACAAAGAGGTGTCATAG 3107

Qy 4126 AGTTTCAAGTTTCAAGCCTGGACACAACAGTCACTCAACAGCAGAGAGCTGTGAAAAGG 4185
Db 3108 AGTTTCAAGTTTCAAGCCTGGACACAACAGTCACTCAACAGCAGAGAGCTGTGAAAAGG 3167
Qy 4186 TCATAGAAAACGGTTGTGATTTTCAGAGACAGGTGAAAAGTTCAGAGTGTGTAGGTGCACACT 4245
Db 3168 TCATAGAAAACGGTTGTGATTTTCAGAGACAGGTGAAAAGTTCAGAGTGTGTAGGTGCACACT 3227
Qy 4246 TATTACCAGCTGAGAAAGTCTCTGCAACGGGTGGCCCACTCTTTCAGCATGCGAGAGG 4305
Db 3228 TATTACCAGCTGAGAAAGTCTCTGCAACGGGTGGCCCACTCTTTCAGCATGCGAGAGG 3287
Qy 4306 ACAGGTACCCCTGGGGCCTGAGTCTCAGCAGAAATCCATCCCAATCATAGTAACTCCTG 4365
Db 3288 ACAGGTACCCCTGGGGCCTGAGTCTCAGCAGAAATCCATCCCAATCATAGTAACTCCTG 3347
Qy 4366 CTCCTCAAAAGCACCTTACATCTCTGACCTACAAGGAGAAATAGCGCATCCAGAGAGAGC 4425
Db 3348 CTCCTCAAAAGCACCTTACATCTCTGACCTACAAGGAGAAATAGCGCATCCAGAGAGAGC 3407
Qy 4426 GATCAGAGGAGAGGACAAAGCCAGATGCTGGTCTCTGATGCTGACGGCAAGGAGATACAG 4485
Db 3408 GATCAGAGGAGAGGACAAAGCCAGATGCTGGTCTCTGATGCTGACGGCAAGGAGATACAG 3467
Qy 4486 CAATCGAAAAGTCTCAAGGCTGAAACCTGAGATCTCTGAACTTTGAGAGTGAAGACAACA 4545
Db 3468 CAATCGAAAAGTCTCAAGGCTGAAACCTGAGATCTCTGAACTTTGAGAGTGAAGACAACA 3527
Qy 4546 AGATTCGTCTGTAACGTCATTTCAGACAGCGTTCACAGTTCGCACGTACAGAACAGCCCC 4605
Db 3528 AGATTCGTCTGTAACGTCATTTCAGACAGCGTTCACAGTTCGCACGTACAGAACAGCCCC 3587
Qy 4606 CCGAAACTCATGCTTATGATTTACAGACCCAGGTTCTGTCATGACAGCTTTGACAGAGGG 4665
Db 3588 CCGAAACTCATGCTTATGATTTACAGACCCAGGTTCTGTCATGCGCTTGGACAGAGGG 3647
Qy 4666 AGCCCAACAGATGCTGGACAAAATGAAAGATGCCAAGATGAAACACCCAGTCCGCGAGC 4725
Db 3648 AGCCCAACAGATGCTGGACAAAATGAAAGTTCGCAAGATGAAACACCCAGTCCGCGAGC 3707
Qy 4726 CCAGAGAGACTTGCAGTTCCTGACCGTTCCTGAGGAGTGGGCTCAGCTTCGGAATGCT 4785
Db 3708 CCAGAGAGACTTGCAGTTCCTGACCGTTCCTGAGGAGTGGGCTCAG - CTCGGAATGCT 3765
Qy 4786 TGCCGCGCTTGCAAGTTCGAAAGCCCGGTGTCAAAGTGAAGATGAAGCTGCTCTCA 4845
Db 3766 TGCCGCGCTTGCAAGTTCGAAAGCCCGGTGTCAAAGTGAAGATGAAGCTGCTCTCA 3825
Qy 4846 ACCCAAGATCCAAAAGGAGCATGCTGCTGATGGCCCTCAGTCCAAAAGCTTAGCCGAGG 4905
Db 3826 ACCCAAGAT - CAAAAGGAGCATGCTGCTGATGGCCCTCAGTCCAAAAGCTTAGCCGAGG 3884
Qy 4906 CAGAGCCAGTGCCTCTGGAACCTTAACCAAGAAATCCCCAGACACCAACGGACCAAGC 4965
Db 3885 CAGAGGCA - - - GTGTCGGAACCTTAACCAAGAAATCCCCAGACACCAACGGACCAAGC 3941
Qy 4966 TAACCCAGAGGAGCGATCCCCCAAAAGTTGAGTTCAGAGAAAGAAATGCTTACCAAGT 5025
Db 3942 TAACCCAGAGGAG - ATGCCCCCAAAAGTTGAGTTCAGAGAAAGAAATGCTTACCAAGT 4000
Qy 5026 CAGTCAAAAGAGAAACAAGGCCCGGACAGAGAGGAGCTTCAGAGGAGGAGGAGGAGCTGG 5085
Db 4001 CAGTCAAAAGAGAAACAAGGCCCGGACAGAGAGGAGCTTCAGAGGAGGAGGAGGAGCTGG 4060
Qy 5086 CAGAAATCTTAAGATGTAGTTGCTGCTTATGATCTGTAAGACCAAGATGTGAAAACAAG 5145
Db 4061 CAGAAATCTTAAGATGTAGTTGCTGCTTATGATCTGTAAGACCAAGATGTGAAAACAAG 4120
Qy 5146 TCACAGAAACAAGATGCTGCTGTTGGGACCTTGAGACCAAGATTTTCAGAGCCCATGATGATC 5205
Db 4121 TCACAGAAACAAGATGCTGCTGTTGGGACCTTG - GACCAAGATTTTCAGAGCCCATGATGATC 4179

Db 421 AAGTCTGAAGAGAGCGCTGCCAACAACCCAGAGGGCAATTGTCAGTGAGGTGGAGAT 480
Qy 1496 GCTGTCTCTCAGGAAAGAAATCAAGGTACAGGGAAGTCCCTTGAAGAAATCTTTCACTAG 1555
Db 481 GCTGTCTCTCAGGAAAGAAATCAAGGTACAGGGAAGTCCCTTGAAGAAATCTTTCACTAG 540
Qy 1556 CTCAAGGCTTAAAGAAAGCTGTCTGGAAAGACAGCAAGGGGAAACAGAGAGGTGGGGAGA 1615
Db 541 CTCAAGGCTTAAAGAAAGCTGTCTGGAAAGACAGCAAGGGGAAACAGAGAGGTGGGGAGA 600
Qy 1616 CGAAGAGCTGGGAATACCAACAATTCACACCGAATCCCCAGAGAGTGTCTGATGAGCA 1675
Db 601 CGAAGAGCTGGGAATACCAACAATTCACACCGAATCCCCAGAGAGTGTCTGATGAGCA 660
Qy 1676 GAAGGGAGAGAGCTCTCGTGTCTCCCGAGAGAGCTTGAGAGACCAAGTGTCTGGAGAA 1735
Db 661 GAAGGGAGAGAGCTCTCGTGTCTCCCGAGAGAGCTTGAGAGACCAAGTGTCTGGAGAA 720
Qy 1736 AGGCGCTGTGAAGCAACCCAGAGATGGGAGCTTGAGAGAGGAACTACTTCCGATGAGCA 1795
Db 721 AGGCGCTGTGAAGCA-CCAGAGATGGGAGAGCTTGAGAGAGGAACTACTTCC- ---GTGGA 775
Qy 1796 GAAGAAGAGAGAGGGATCACTCCCTGGGCATCTCTCAAAAAGATGGTGACACCCCAAGAA 1855
Db 776 GAGAAGAGAGGAGGATCACTCCCTGGGCATCTCTCAAAAAGATGGTGACACCCCAAGAA 835
Qy 1856 ACGGCTCCGAAGACCTTTCTGAGAGTGACAAGGAGAAAGAGCTGGAGAGGTCAAGAGCGC 1915
Db 836 AC-GGTCCGAAGACCTTTCTGAGAGTGACAAGGAGAAAGAGCTGGAGAGGTCAAGAGCGC 894
Qy 1916 CACTTGTCTCTCACTATAGCACAGTGTGAGAAATCCAGATGAAGTCAAACTGTGTGG 1975
Db 895 CACTTGTCTCTCACTATAGCACAGTGTGAGAAATCCAGATGAAGTCAAACTGTGTGG 954
Qy 1976 TGAGGAACAAAGCCAGAGAAACCAAGCGTGTAGGTCGATCTTCACTGTCTTGGGAAGC 2035
Db 955 TGAGGAACAAAGCCAGAGAAACCAAGCGTGTAGGTCGATCTTCACTGTCTTGGGAAGC 1014
Qy 2036 ACTGATTTGTGTGCGATCATCCAGAAAGAGAGCAAGAAAGGCACTCTTTCAGATGATGA 2095
Db 1015 ACTGATTTGTGTGCGATCATCCAGAAAGAGAGCAAGAAAGGCACTCTTTC- ---GATAT 1071
Qy 2096 AGGAGGCGCAAGACACTGGGA- GGGACACAGTACAGAGCAGAGGAGCCAGCAAGACAA 2154
Db 1072 AAGAGGCGCAAGACACTGGGAGGGGACAGTACAGAGCAGAGGAGCCAGCAAGACAA 1131
Qy 2155 AAGAAGCGGAAACAGACGCTGTCTCCAGACCCAGGAGCAGGACCAAGCGCAAGGAA 2214
Db 1132 AAGAGCC- GAAACAGAGCTGTCTCTCCAGACCCAGGAGCAGGACCAAGCGCAAGGAA 1190
Qy 2215 GTTCTCACCAGCCAGCGGAGAGCCCTTCCGAAGGGGAAGGTGTCTCCACTTTGGGAGT 2274
Db 1191 GTTCTCACCAGCCAGCGGAGAGCCCTTCCGAAGGGGAAGGTGTCTCCACTTTGGGAGT 1250
Qy 2275 CATTTAAAGATTAGTCACTCCAGAAAGAAATTCAGAGTCAAACTGGAAGAGAAAGCCG 2334
Db 1251 CATTTAAAGATTAGTCACTCCAGAAAGAAATTCAGAGTCAAACTGGAAGAGAAAG 1310
Qy 2335 AAGACTCTAGTGTAG- ---AGCAGTTGTCCACTGAGATCGAACCGAGTGAAGAGA 2386
Db 1311 CCGAAGGACTCTAGTGTAGGAGCAGGTTGTCCACTGAGATCGAACCGGTTGTGAAGAGA 1370
Qy 2387 ATCTTGGGTTTCCATTAAAGAAATTCATCCCGGACGGGAGAAAGAGGCGAGACGGGAA 2446
Db 1371 ATCTTGGGTTTCCATTAAAGAAATTCATCCCGGACGGGAGAAAGAGAGGCGAGATGGGAA 1430
Qy 2447 -GCAAGAACAGCCACTGTGGAGAGACTCAGGGCCAGTGGAGATAAATGAGGACGACCTTA 2505
Db 1431 GGCAAGAACAGCCACTGTGGAGAGACTCAGGGCCAGTGGAGATAAATGAGGACGAGCCTG 1490
Qy 2506 ATGTCCAGCGCTGTGCTCTGTCTGAGTATATCCAGTGGAGGGGAGAGATGGAAG 2565
Db 1491 ATGTCCAGCGCTGTGCTCTGTCTGAGTATGATGCAGTGGGAGGGGAGAGATGGAAG 1550

Qy 2566 CCAGGGGAATACGGAGCTGCCCCAGCTGCTGGGGCTGTGTACGTGTCCGAGGAGCTCA 2625
Db 1551 CCAGGGGAATACGGAGCTGCCCCAGCTGCTGGGGCTGTGT- ---AGTGTCCGAGGAGCTCA 1607
Qy 2626 GTAAAGACTCTGGTCCACACTGTGAGTGTCCAGTCAATTGATGGGACCGAGGCGAGTCA 2685
Db 1608 GTAAAGACTCTGGTCCACACTGTGAGTGTCCAGTCAATTGATGGGACCGAGGCGAGTCA 1667
Qy 2686 GTGTGAAAGAGCGGTCTCTCTTCGTGGATATCCCGTTCCGTAACAGAACCTTTTGAACACA 2745
Db 1668 GTGTGAAAGAGCGGTCTCTCTTCGTGGATATCCCGTTCCGTAACAGAACCTTTTGAACACA 1727
Qy 2746 CAGCGGAGAGCCATGCCACCTGTTTGAAGAGGTCACTGAAAAGACATCAITTCGAGAAG 2805
Db 1728 CAGCGGAGAGCCATGCCACCTGTTTGAAGAGGTCACTGAAAAGACATCAITTCGAGAAG 1787
Qy 2806 AAACCTCTGTGCTCACCCAGAGCTTACCCAGAGGTTAAAGTGCCTCATGACACATGGTCA 2865
Db 1788 AAACCTCTGTGCTCACCCAGAGCTTACCCAGAGGTTAAAGTGCCTCATGACACATGGTCA 1847
Qy 2866 CCAGTGAAGTGAATTTACCTCAGAAAGCTGTGACAGCACAAGAGACCTCAGAGGCTCTCC 2925
Db 1848 CCAGTGAAGTGAATTTACCTCAGAAAGCTGTGACAGCACAAGAGACCTCAGAGGCTCTCC 1907
Qy 2926 GTACTGAAGAGATTACCGAAGCATCGGGGGCGAAGAGACCAAGAGACATGGTGTCCGCGAG 2985
Db 1908 GTACTGAAGAGATTACCGAAGCATCGGGGGCGAAGAGACCAAGAGACATGGTGTCCGCGAG 1967
Qy 2986 TTTCCAGCTGACTGACTCTCCAGACACACAGAGGAAGCCACCCAGAGGTTCAGAGGTTAG 3045
Db 1968 TTTCCAGCTGACTGACTCTCCAGACACACAGAGGAAGCCACCCAGAGGTTCAGAGGTTAG 2027
Qy 3046 AGAGTGTGTGTGTAGATACAGAAAGAGAGAGCGCCAGAGCGAGGCGCATCTCCAGAGCG 3105
Db 2028 AGGTTGTGTGTGTAGATACAGAAAGAGAGAGCGCCAGAGCGAGGCGCATCTCCAGAGCG 2087
Qy 3106 TTGCAGACAAGGTGAAAGAGGAGTCCAGGTGCTCTCAACCCAGAGCTGTGCAGAGAACGG 3165
Db 2088 TTGCAGACAAGGTGAAAGAGGAGTCCAGGTGCTCTCAACCCAGAGCTGTGCAGAGAACGG 2147
Qy 3166 GGTCAAAAGCACTGGAGAGAGTTGAGGAGGTAGAGAGAGACTCCGAAGTGTGCTGCTCGG 3225
Db 2148 GGTCAAAAGCACTGGAGAGAGTTGAGGAGGTAGAGAGGAGTCCGAAGTGTGCTGCTCGG 2207
Qy 3226 AGAAAGAGAGAGAGCTTATGCCGAAAGGACCCGTGCAGAGAGCTGAGAGCTGAGCATCTTG 3285
Db 2208 AGAAAGAGAGAGAGCTTATGCCGAAAGGACCCGTGCAGAGAGCTGAGAGCTGAGCATCTTG 2267
Qy 3286 CACAGGCTCTGAGACTGGACAGGCTTACTCCAGAGAGCCCTTGAAAGTTCCTGAAAGTCA 3345
Db 2268 CACAGGCTCTGAGACTGGACAGGCTTACTCCAGAGAGCCCTTGAAAGTTCCTGAAAGTCA 3327
Qy 3346 CAGATGTAGACCAATGTGCGCACGTGCCAGGTTATCAAGTCCAGAGAGCTGATGGAAACAG 3405
Db 2328 CAGATGTAGACCAATGTGCGCACGTGCCAGGTTATCAAGTCCAGAGAGCTGATGGAAACAG 2387
Qy 3406 CCGTGGCCCTCGAGTCACTCCGAAACCTTGACAGACAGTGAACAAATGGAAGCACTCCCT 3465
Db 2388 CCGTGGCCCTCGAGTCACTCCGAAACCTTGACAGACAGTGAACAAATGGAAGCACTCCCT 2447
Qy 3466 TAGCAGATTTCAGACACTGCAGATGGGACACAGCAAGATGAAACCAATTGACAGCAGGACA 3525
Db 2448 TAGCAGATTTCAGACACTGCAGATGGGACACAGCAAGATGAAACCAATTGACAGCAGGACA 2507
Qy 3526 GTAAAGCCACTCGAGTGTGAGGCACTCAGGTCAAGAGAGAGGCGGCTACTGCTC 3585
Db 2508 GTAAAGCCACTCGAGTGTGAGGCACTCAGGTCAAGAGAGAGGCGGCTACTGCTC 2567
Qy 3586 AGAAAGAGAGGCTTCGACACTACCTTAATAATGTTCCAGCCCGAGGAGAAACATGGGGAAG 3645
Db 2568 AGAAAGAGAGGCTTCGACACTACCTTAATAATGTTCCAGCCCGAGGAGAAACATGGGGAAG 2627

QY	3646	AACCAAGAAAGATGTTCTTGAACCTACACAGAAAGCTTACTGTGACGCGTCCCG	3705
Db	2628	AACCAAGAAAGATGTTCTTGAACCTACACAAAGAGCTTCTGTGACGCGTCCCG	2687
QY	3706	TTTCGGCAAAAGCTGAGGGGGGTCAAGAGGGTGAGGTTGACTGGTTGATGAGAAAAAG	3765
Db	2688	TTGCGCAAAAGCTGAGGGGGGTCAAGAGGGTGAGGTTGACTGGTTGATGAGAAAAAG	2747
QY	3766	TCAAAAGAAACAGAGAGGTGTTTGTACACTTGACCCCAACAGTCAAAAGCTGTGATG	3825
Db	2748	TCAAAAGAAACAGAGAGGTGTTTGTACACTTGACCCCAACAGTCAAAAGCTGTGATG	2807
QY	3826	TGACATATGACAGTGAATGATGGGAGTGGCGGGGTGTCAAGAAAAAGAGATGCTGAAG	3885
Db	2808	TGACATATGACAGTGAATGATGGGAGTGGCGGGGTGTCAAGAAAAAGAGATGCTGAAG	2867
QY	3886	TGCACAGTTTGTAGCCCTGGAGGAGGGGAGATGGAAACTGACGTTGAAAAAGAAAGG	3945
Db	2868	TGCACAGTTTGTAGCCCTGGAGGAGGGGAGATGGAAACTGACGTTGAAAAAGAAAGG	2927
QY	3946	AGACAAAGCCACAGCAAGTAGAGTGAAGAGGTGAGCAGGAAACAGCGCTCTGAGCATG	4005
Db	2928	AGACAAAGCCACAGCAAGTAGAGTGAAGAGGTGAGCAGGAAACAGCGCGTCTGAGCATG	2987
QY	4006	AAGAAACCTACGGGAACCAAGTCTTGACACTTGACATGCCCACTCAAGAGGGGGAGG	4065
Db	2988	AAGAAACCTACGGGAACCAAGTCTTGACACTTGACATGCCCACTCAAGAGGGGGAGG	3047
QY	4066	CACGTGGGAAGCCTTGGAGAGAGCCCTTCTCCACAGCCAAAGCAAGCGTTGCTAG	4125
Db	3048	CACGTGGGAAGCCTTGGAGAGAGCCCTTCTCCACAGCCAAAGCAAGCGTTGCTAG	3107
QY	4126	AGGTTCAAGTTCAAAGCCTGGAACAACACTCACTCAACAGAGAAAGCTGTGAAAAGG	4185
Db	3108	AGGTTCAAGTTCAAAGCCTGGAACAACAACACTCACTCAACAGAGAAAGCTGTGAAAAGG	3167
QY	4186	TCATAGAAACGGTGTGATTTCAAGACAGGTGAAGTCAAGAGTGTATAGTGCACACT	4245
Db	3168	TCATAGAAACGGTGTGATTTCAAGACAGGTGAAGTCAAGAGTGTATAGTGCACACT	3227
QY	4246	TATTACACAGTGAAGTCCCTGCAACGGGTGGCCACTGACACTTCACATGACAGAG	4305
Db	3228	TATTACACAGTGAAGTCCCTGCAACGGGTGGCCACTGACACTTCACATGACAGAG	3287
QY	4306	ACAACGTTACCCCTGGGGCCTGAAGTCTCAGGCAGAAATCCATCCAAATAGTAACCTCTG	4365
Db	3288	ACAACGTTACCCCTGGGGCCTGAAGTCTCAGGCAGAAATCCATCCAAATAGTAACCTCTG	3347
QY	4366	CTTCCTGAAAGCACCTTACATCTTGACCTCAAGGAAATTAAGGCATCCCAAGAGAGC	4425
Db	3348	CTTCCTGAAAGCACCTTACATCTTGACCTCAAGGAAATTAAGGCATCCCAAGAGAGC	3407
QY	4426	GATCAGAGGAAGAGACAACCCAGATGCTGTCTGTATGCTGAACGGGAGAGATGACAG	4485
Db	3408	GATCAGAGGAAGAGACAACCCAGATGCTGTCTGTATGCTGAACGGGAGAGATGACAG	3467
QY	4486	CAATCGAAAAAGTCTCAAGGCTGAACCTGAGATCTGTGAACCTTGAGAGTAAAGCAACA	4545
Db	3468	CAATCGAAAAAGTCTCAAGGCTGAACCTGAGATCTGTGAACCTTGAGAGTAAAGCAACA	3527
QY	4546	AGATTGTCTGAACGTCATTCAAGACCGGTGACCAAGTTGCAAGTGAACAGAAACAGCCC	4605
Db	3528	AGATTGTCTGAACGTCATTCAAGACCGGTGACCAAGTTGCAAGTGAACAGAAACAGCCC	3587
QY	4606	CCGAAACTCATGCTTATGATTTCAACAGCCAGAGTTCCTGCATGCAAGGCTTGACAGCAGG	4665
Db	3588	CCGAAACTCATGCTTATGATTTCAACAGCCAGAGTTCCTGCATGCAAGGCTTGACAGCAGG	3647
QY	4666	AGCCCAACAGATGCTGACAAAAATGAAGAATGCAAGATGAACACCACTGCGCAGC	4725
Db	3648	AGCCCAACAGATGCTGACAAAAATGAAGAATGCAAGATGAACACCACTGCGCAGC	3707
QY	4726	CCAGAGAGGACTTGCAAGTCTGAACCGTCTGAGAGCAATGGGCTCAAGCTTCGAAAAATGCT	4785

Db	3708	CGAGAGGACCTTGCAAGTCCGACCGCTTCGAGGCAATGGCTCAG--CTCCGAAATGCT	3765
QY	4786	TGCCGCGCTTGCAAGTTGAAAGCGCCGGGTGTCAAGTAAAGCAATTGAGAACTGCTCTCTCA	4845
Db	3766	TGCCGCGCTTGCAAGTTGAAAGCGCCGGGTGTCAAGTAAAGCAATTGAGAAAGCTGCTCTCTCA	3825
QY	4846	ACCCAAAGATCCAAAAGAGCAATGCTGTATGAGGCGCTCAAGCTCAAAAGCTTGAAGCCCAAG	4905
Db	3826	ACCCAAAGAT-CAAAAAGAGCAATGCTGTATGAGGCGCTCAAGCTCAAAAGCTTGAAGCCCAAG	3884
QY	4906	CAGAGCCAGTGCCTCTGAGAACTTAAACCAAGAAATCCCGACCAACCAACGACCAAGC	4965
Db	3885	CAGAGGCA---GTGTCTGGAACCTTAACCAAGAAATCCCGACCAACCAACGACCAAGC	3941
QY	4966	TAAACCGAGAGGGCGATCCCCCAAAAGTTGAGGTCCAGAGAAAGAAATGTCTAACCAAGT	5025
Db	3942	TAAACCGAGAGGCG-ATGCCCCCAAAAGTTGAGGTCCAGAGAAAGAAATGTCTAACCAAGT	4000
QY	5026	CAGTCAAAGAGAAACAAGGCCCGAGAGAGAGACCTGACAGACCCAAAGAGAGACCTGG	5085
Db	4001	CAGTCAAAGAGAAACAAGGCCCGAGAGAGAGACCTGACAGACCCAAAGAGAGACCTGG	4060
QY	5086	CAGAACTCTTAAGATGTTAGTGTCTCATGTGCATCTGTAAACAGACAGATGTGAAAAACAAG	5145
Db	4061	CAGAACTCTTAAGATGTTAGTGTCTCATGTGTCTATCTGTAAACAGACAGATGTGAAAAACAAG	4120
QY	5146	TCACAGAAACAAGATGCTGCTGTTGGAGACCTTGAGACCAAGATTTTCAGAGCCCATGAGATC	5205
Db	4121	TCACAGAAACAAGATGCTGCTGTTGGAGACCTTG-AGCAAGATTTTCAGAGCCCATGAGATC	4179
QY	5206	CAGAGACACAGAGCCCGTGTCCAAATGATTTTCAACCCAGTAGAGAGCACCCCGACATTCAGAGCT	5265
Db	4180	CAGAGACACAGAGCCCGTGTCCAAATGATTTTCAACCCAGTAGAGAGCACCCCGAATTCAGAGCT	4239
QY	5266	TCATCGGGAAGCTTAGAGCAGCAGCTAACATTTCTCGTTTCAAGACTGCTTGATTTGGCCC	5325
Db	4240	TCATCGGGAAGCTTAGAGCAGCAGCTAACATTTCTCGTTTCAAGACTGCTTGATTTGGCCC	4299
QY	5326	TTGATGCCGTCCTGTATTTCTAACATTTTCTCGTTTCAAGACTGCTTGATTTGGCCC	5385
Db	4300	TTGATGCCGTCCTGTATTTCTAACATTTTCTCGTTTCAAGACTGCTTGATTTGGCCC	4359
QY	5386	TTGATGCCGTCCTGTATTTTCCGATTTAAGGTCCMGCGCTTCAACCTGGAACCAATTC	5445
Db	4360	TTGATGCCGTCCTGTATTTTCCGATTTAAGGTCCMGCGCTTCAACCTGGAACCAATTC	4419
QY	5446	GCAATACCTAGTTTCAACTTCTCAAACTGAGAGACTCCTCTTATGTATTTATATGATGT	5505
Db	4420	GCAATACCTAGTTTCAACTTCTCAAACTGAGAGACTCCTCTTATGTATTTATATGATGT	4479
QY	5506	TTTATGTAGTCTCTCTGTACCTATGTATATTTTTTTCTAACGTTTAAAGCAATGCT	5565
Db	4480	TTTATGTAGTCTCTCTGTACCTATGTATATTTTTTTCTAACGTTTAAAGCAATGCT	4539
QY	5566	TTTTGTATTTATGCAATATTAAGGGGTGACGCCATAGACCGCTTTGAAAACCTCCAA	5625
Db	4540	TTTTGTATTTATGCAATATTAAGGGGTGACGCCATAGACCGCTTTGAAAACCTCCAA	4599
QY	5626	GCCCTCACTGTAACTTCCAGCAAAACAGATTAACATTCCTGAGCAAGAGAGCAAGCTTTTT	5685
Db	4600	GCCCTCACTGTAACTTCCAGCAAAACAGATTAACATTCCTGAGCAAGAGAGCAAGCTTTTT	4659
QY	5686	TTTAAATTTACTGATGCTTAGATCTGTGGGCTTCTAGTCTCTGAAAAGTGTTGTTTTCC	5745
Db	4660	TTTAAATTTACTGATGCTTAGATCTGTGGGCTTCTAGTCTCTGAAAAGTGTTGTTTTCC	4719
QY	5746	TATGACACGCAAGCTCAGAAATATAAAACCCCAATTTTGAACAATCCAGATGTCCCAATAT	5805
Db	4720	TATGACACGCAAGCTCAGAAATATAAAACCCCAATTTTGAACAATCCAGATGTCCCAATAT	4779
QY	5806	TACCAATTTTTTCCCCCTTTTGTCTATCAGTCCAGGTGTGAAAAGAAAGTCTCTCT	5865

```
Db 4780 TACCATGATTTTTCCTCCCTTTTTCCTTAATCCAGTCCAGGTTGGAAGAAGTCTCTCT 4839
Qy 5866 GTGTGAGATTAAAGCCCTGTCTCTTAATGATGATGACAAATGAGTGTGCTTAAGGCCATGA 5925
Db 4840 GTGTGAGATTAAAGCCCTGTCTCTTAATGATGATGACAAATGAGTGTGCTTAAGGCCATGA 4899
Qy 5926 GATCTTTTCTTAATCGAGAGAAATCTGTTGTACGTTTTTTTGTGATGTACTCTTCTATGCT 5985
Db 4900 GATGTTTCTTAATCGAGAGAAATCTGTTGTACGTTTTTTTGTGATGTACTCTTCTATGCT 4959
Qy 5986 GGACCGCAATTCATATGATGATGAGTGAAGTGAAGTCCCTGTTCTTTTACAGATGGTATTTTGTATG 6045
Db 4960 GGACCGCAATTCATATGATGATGAGTGAAGTGAAGTCCCTGTTCTTTACAGATGGTATTTTGTATG 5019
Qy 6046 ATATCGGAGTTTGTCTGTGTTATATCTGTGCCCCCTTTTAAAGAACAAATGTTGCAATTATG 6105
Db 5020 ATATCGGAGTTTGTCTGTGTTATATCTGTGCCCCCTTTTAAAGAACAAATGTTGCAATTATG 5079
Qy 6106 TTCTTTTGGATAAATTTGATTTTCACTCACTGATTTTAAATTAACATATTTGACTAC 6160
Db 5080 TTCTTTTGGATAAATTTGATTTTCACTCACTGATTTTAAATTAACATATTTGACTAC 5134

RESULT 9
US-09-902-432-1
; Sequence 1, Application US/09902432
; GENERAL INFORMATION:
; APPLICANT: Susan G. Jaken
; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
; FILE REFERENCE: A30558-A-FWC-A, 070156.0597
; CURRENT APPLICATION NUMBER: US/09/902.432
; CURRENT FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 08/978,277
; PRIOR FILING DATE: 1997-11-25
; PRIOR APPLICATION NUMBER: 08/665,401
; PRIOR FILING DATE: 1996-06-18
; PRIOR APPLICATION NUMBER: 08/635,121
; PRIOR FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 5134
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-09-902-432-1

Query Match 78.48; Score 4831.4; DB 34; Length 5134;
Best Local Similarity 97.98; Pred. No. 0;
Matches 5048; Conservative 0; Mismatches 76; Indels 31; Gaps 14;

Qy 1016 GGAAGAAGACAGAGCCAGCTCGGAGGAGCAGGAGCCGGCAGAAAGACACAGACCAGGCCAG 1075
Db 1 GGAAGAAGACAGAGCCAGCTCGGAGGAGCAGGAGCCGGCAGAAAGACACAGACCAGGCCAG 60

Qy 1076 GTTGTGACGACATACGAGAAGTGGAGCTGCTTTTGAAGACACAGGTTGTTGACCTGGA 1135
Db 61 GTTGTGACGACATACGAGAAGTGGAGCTGCTTTTGAAGACACAGGTTGTTGACCTGGA 120

Qy 1136 GGCATGTCAGAGAGAAGTGTCTCTTTGGCAACGGAAGTGTGATGAGAAGATGGA 1195
Db 121 GGCATGTCAGAGAGAAGTGTCTCTTTGGCAACGGAAGTGTGATGAGAAGATGGA 180

Qy 1196 AGCCCAACCAAGAAGTTGTTGACAGAGTCCAGTGCAGACCGTGGAGAGACAGAGGAGGA 1255
Db 181 AGCCCAACCAAGAAGTTGTTGACAGAGTCCAGTGCAGACCGTGGAGAGACAGAGGAGGA 240

Qy 1256 GCAGGGAGGAGGAGAGGCTGAAGGGGGCTGTGTGTAGAAAGAAACAGGAGAAATCCTT 1315
Db 241 GCAGGGAGGAGGAGAGGCTGAAGGGGGCTGTGTGTAGAAAGAAACAGGAGAAATCCTT 300

Qy 1316 GCCCCTTGAGAAATCGGCTGAGCCCGAGGAGGTCCTCCCGAGGAGCTGAGCTGCTGAGGA 1375
```

```
Db 301 GCCCCTTGAGAAACTGGCTGAGCCCCAGGAGGTCCTCCAGGAAGCTGAGCCTGCTGAGGA 360
Qy 1376 GCTGATGAAGACAGCAGAGAGATGTGTCTCTGGAGGAGACCACTCAACTGACAGACCT 1435
Db 361 GCTGATGAAGACAGCAGAGAGATGTGTCTCTGGAGGAGACCACTCAACTGACAGACCT 420
Qy 1436 AAGTCTTGAAGAGAAGCGCTGCCAAAACACCCAGAAAGGCAATTTGTAGTGAAGTGGAGAT 1495
Db 421 AAGTCTTGAAGAGAAGCGCTGCCAAAACACCCAGAAAGGCAATTTGTAGTGAAGTGGAGAT 480
Qy 1496 GCTGTCTCTCTCAGGAAGAATCAAGGTACAGGGAAGTCCCTTGAAGAAAATCTTTCACTAG 1555
Db 481 GCTGTCTCTCTCAGGAAGAATCAAGGTACAGGGAAGTCCCTTGAAGAAAATCTTTCACTAG 540
Qy 1556 CTCAAGGCTTAAAGAAGCTGTCTGGGAAGAACAGAGAGGGGAAAACGAGAGGTCGGGGAGA 1615
Db 541 CTCAAGGCTTAAAGAAGCTGTCTGGGAAGAACAGAGAGGGGAAAACGAGAGGTCGGGGAGA 600
Qy 1616 CGAAGAGCCTGGAGATACCAACACATTTACACCGAATCCCCAGAGAGTCTGATGAGCA 1675
Db 601 CGAAGAGCCTGGAGATACCAACACATTTACACCGAATCCCCAGAGAGTCTGATGAGCA 660
Qy 1676 GAAGGGAGAGAGCTCTGCGTCTGCTCCCGAGGAGCCTGAGGAGACCACTGCTTGGAGAA 1735
Db 661 GAAGGGAGAGAGCTCTGCGTCTGCTCCCGAGGAGCCTGAGGAGACCACTGCTTGGAGAA 720
Qy 1736 AGGGCCGCTCGAAGCACTGAGGAGTGGGAAGCTGAGGAAGGAATCTTCCGATGAGA 1795
Db 721 AGGGCCGCTCGAAGCACTGAGGAGTGGGAAGCTGAGGAAGGAATCTTCCGATGAGA 775
Qy 1796 GAAGAAGAGAGAGGATCACTCCTGGGCACTCTTCAAAAAGATGGTGAACCCAAAGAA 1855
Db 776 GAGAAGAAGAGAGGATCACTCCTGGGCACTCTTCAAAAAGATGGTGAACCCAAAGAA 835
Qy 1856 ACGGTCCGAGAGACCTTCTGAGAGTGCACAGGAGGAGAGCTGGAAGAGTCAAGAGCGC 1915
Db 836 AC-GGTCCGAGAGACCTTCTGAGAGTGCACAGGAGGAGAGCTGGAAGAGTCAAGAGCGC 894
Qy 1916 CACTTGTCTCTCCACTGATAGCAGTGTGAGAAATGCAAGATGAACTCAAACTGTG 1975
Db 895 CACTTGTCTCTCCACTGATAGCAGTGTGAGAAATGCAAGATGAACTCAAACTGTG 954
Qy 1976 TGAGGAACAAAAGCCAGAGAAACAAAGCTAGGGTGGATATCTTCACTGCTTTGGGAAGC 2035
Db 955 TGAGGAACAAAAGCCAGAGAAACAAAGCTAGGGTGGATATCTTCACTGCTTTGGGAAGC 1014
Qy 2036 ACTGATTTGTGTCGATCATCCAAAGAGAGACCAAGAGGCAATCTTCTCAGATGATGA 2095
Db 1015 ACTGATTTGTGTCGATCATCCAAAGAGAGACCAAGAGGCAATCTTCTCAGATGATGA 1071
Qy 2096 AGGAGGCCCAAGGACACTGGGA-GGGGACAGTCAAGAGAGAGGAGGCCAGCAAGACA 2154
Db 1072 AAGAGGCCCAAGGACACTGGAGGGGGACAGTCAAGAGAGAGGAGGCCAGCAAGACA 1131
Qy 2155 AAGAAGCCGGAACAGAGCTGTTCTGCCAGCACCCAGAGAGGAGCCAAAGCGCAAGAA 2214
Db 1132 AAGAAGCC-GAAACAGAGCTGTTCTGCCAGCACCCAGAGAGGAGCCAAAGCGCAAGAA 1190
Qy 2215 GTTCTCACCCGAGCCAGCGGGAAGCCCTTCCGAAGGGGAGGTTCTTCCACTTTGGAGT 2274
Db 1191 GTTCTCACCCGAGCCAGCGGGAAGCCCTTCCGAAGGGGAGGTTCTTCCACTTTGGAGT 1250
Qy 2275 CATTTAAAGATTAGTCACTCCCAAGAAAAAATCCAAGTCAAAATCTGGAAGAGAAAGCCG 2334
Db 1251 CATTTAAAGATTAGTCACTCCCAAGAAAAAATCCAAGTCAAAATCTGGAAGAGAAAG 1310
Qy 2335 AAGACTCTAGTGTAG-----AGCAGTTGTCTCACTGAGATCGAACCCGATGAGAGA 2386
Db 1311 CCGGAAGGACTCTAGTTGTAGGAGCAGGTTGTCCACTGAGATCGAACCCGTTAGAGAAGA 1370
Qy 2387 ATCTTGGGTTTCCATTTAAGAAATTCATCCCCGAGCGCGGGAAGAAAGGGGAGCGGAA 2446
Db 1371 ATCTTGGGTTTCCATTTAAGAAATTCATCCCCGAGCGCGGGAAGAAAGGGGAGCGGAA 1430
```

QY 2447 -GCAAGAACAGCCACTGTGAGAGCTCAGGGCCAGTGGAGTAAATGAGACACCTTA 2505
Db 1431 GGCAGAAACAAGCCACTGTGGAGAGCTCAGGGCCAGTGGAGTAAATAGAGCGAGCTTG 1490
QY 2506 ATGTCCCGCCGTGTGCTGTGTGTGATTAATGACGTGGAGGGGAGAAATGTAAG 2565
Db 1491 ATGTCCCGCCGTGTGCTGTGTGTGATTAATGACGTGGAGGGGAGAAATGTAAG 1550
QY 2566 CCGAGGGGAATACGAGAGCTGCCCGCAGCTGTGGGGCTGTGTACGTGTCCGAGAGCTCA 2625
Db 1551 CCGAGGGGAATACGAGAGCTGCCCGCAGCTGTGGGGCTGTGTACGTGTCCGAGAGCTCA 1607
QY 2626 GTAAGACTCTGTGTCCACACTGTGAGTGTGCGAGTATGATGGAGCAAGGGACAGTCCACA 1667
Db 1608 GTAAGACTCTGTGTCCACACTGTGAGTGTGCGAGTATGATGGAGCAAGGGACAGTCCACA 1667
QY 2666 GTGTGAGAGAGGGGTCTCTGTGTGTGATATCCGCTTCGGTAAACAGACCTTTGAACACA 2745
Db 1668 GTGTGAGAGAGGGGTCTCTGTGTGTGATATCCGCTTCGGTAAACAGACCTTTGAACACA 1727
QY 2746 CAGCGGAGAGAGCCATGCGACCTGTGTAAGAGTCACTGAATAAGACATCATTTGAGAG 2805
Db 1728 CAGCGGAGAGAGCCATGCGACCTGTGTAAGAGTCACTGAATAAGACATCATTTGAGAG 1787
QY 2806 AAATCTCTGTGTCAACCAAGAGTTCACAGAGGTTAAAGTCCCATGACAGATGTCA 2865
Db 1788 AAATCTCTGTGTCAACCAAGAGTTCACAGAGGTTAAAGTCCCATGACAGATGTCA 1847
QY 2866 CCAGTGAAGTGAATTTCACTTCAGAGCTGTGACAGCCAGAGAGCTTCAGAGGCTCTCC 2925
Db 1848 CCAGTGAAGTGAATTTCACTTCAGAGCTGTGACAGCCAGAGAGCTTCAGAGGCTCTCC 1907
QY 2926 GTACTGAAGAAATTACCAAGAGATCGGGGGCCGAAGAGCACAGACATGTGTCCGAG 2985
Db 1908 GTACTGAAGAAATTACCAAGAGATCGGGGGCCGAAGAGCACAGACATGTGTCCGAG 1967
QY 2986 TTTTCCAGCTGACTGACTCTCCAGACACCAAGAGAGAGCCCAAGTTTCAGAGGTAG 3045
Db 1968 TTTTCCAGCTGACTGACTCTCCAGACACCAAGAGAGAGCCCAAGTTTCAGAGGTAG 2027
QY 3046 AGAGTGTGTGTATATATCAGAGAGAGAGAGGAGCGCAGAGCCATCCCTCAAGCG 3105
Db 2028 AGAGTGTGTGTATATATCAGAGAGAGAGAGGAGCGCAGAGCCATCCCTCAAGCG 2087
QY 3106 TTGCAGACAGAGTGAAGAGAGAGTCCAGAGTGCCTGCAACCCAGACTGTGACAGAGAG 3165
Db 2088 TTGCAGACAGAGTGAAGAGAGAGTCCAGAGTGCCTGCAACCCAGACTGTGACAGAGAG 2147
QY 3166 GGTCAAAAGCACTGGAGAGAGTTGAGAGAGTGAAGAGAGTCCGAAGTGTGCTTGG 3225
Db 2148 GGTCAAAAGCACTGGAGAGAGTTGAGAGAGTGAAGAGAGTCCGAAGTGTGCTTGG 2207
QY 3226 AGAAGAGAGAGAGCTTTATGCGGAAGAGCCGTCAGAGAGCTGGAAGTGTGACATTTG 3285
Db 2208 AGAAGAGAGAGAGCTTTATGCGGAAGAGCCGTCAGAGAGCTGGAAGTGTGACATTTG 2267
QY 3286 CACAGGGCTCTGAGACTGGAACAGGCTACTCAGAGAGCTTGAAGTCTTGAAGTCAAG 3345
Db 2268 CACAGGGCTCTGAGACTGGAACAGGCTACTCAGAGAGCTTGAAGTCTTGAAGTCAAG 2327
QY 3346 CAGATGTAGACCATGTGTGCGCAGTGTCCAGGTTATCAAGTCTCAGAGCTGATGGAACAG 3405
Db 2348 CAGATGTAGACCATGTGTGCGCAGTGTCCAGGTTATCAAGTCTCAGAGCTGATGGAACAG 2387
QY 3406 CCGTGGCCCTTGAATCTTCCGAAACCTTGAACAGATGAGCAAAATGGAACATCTCCCT 3465
Db 2388 CCGTGGCCCTTGAATCTTCCGAAACCTTGAACAGATGAGCAAAATGGAACATCTCCCT 2447
QY 3466 TAGCAGATTCAAGACTGAGATGGAGACAGCAAGATGAAACATTGACAGCCAGAGACA 3525
Db 2448 TAGCAGATTCAAGACTGAGATGGAGACAGCAAGATGAAACATTGACAGCCAGAGACA 2507

QY 3526 GTAAAGCACTGCAAGCTGTGAGAGGACAGAGTCAAGAGAGAGAGCGGCTACTGTCTC 3585
Db 2508 GTAAAGCACTGCAAGCTGTGAGAGGACAGAGTCAAGAGAGAGAGCGGCTACTGTCTC 2567
QY 3586 AGAAGAGAGAGCCCTTGACACTACTAATATATGTTCCAGGCCAGAGAGAAACATGGGAG 3645
Db 2568 AGAAGAGAGAGCCCTTGACACTACTAATATATGTTCCAGGCCAGAGAGAAACATGGGAG 2627
QY 3646 AACCAAGAGAGATGTTCTTGAACCTTACACAGACAGAGCTTACTGTGACAGCCGTCCG 3705
Db 2628 AACCAAGAGAGATGTTCTTGAACCTTACACAGACAGAGCTTACTGTGACAGCCGTCCG 2687
QY 3706 TTTGCGAAAAGACTGAGAGTGGTCAAGAGGTGAGTTGACTGGTTGGATGGAGAAAAG 3765
Db 2688 TTTGCGAAAAGACTGAGAGTGGTCAAGAGGTGAGTTGACTGGTTGGATGGAGAAAAG 2747
QY 3766 TCAAGAGAGAGACAGAGAGTGTGTTGTACACTGTGACCTGAGCCCAACAGTCAAAAGCTGTGATG 3825
Db 2748 TCAAGAGAGAGACAGAGAGTGTGTTGTACACTGTGACCTGAGCCCAACAGTCAAAAGCTGTGATG 2807
QY 3826 TGACATATGACATGTAAGTGAATGGAGTGGCGGGGTGTCAAGAAAAGAGAGTACTGAAG 3885
Db 2808 TGACATATGACATGTAAGTGAATGGAGTGGCGGGGTGTCAAGAAAAGAGAGTACTGAAG 2867
QY 3886 TGCAGAGCTTTAGGCTGTGAGAGGAGAGAGATGGAACGTGACGTTGAAGAGAGAAAAGGG 3945
Db 2868 TGCAGAGCTTTAGGCTGTGAGAGGAGAGAGATGGAACGTGACGTTGAAGAGAGAAAAGGG 2927
QY 3946 AGACAAAGCCAGAGCAAGTGAAGTGAAGAGGTGAGACAGAGAAACAGCCGCTTGAAGATG 4005
Db 2928 AGACAAAGCCAGAGCAAGTGAAGTGAAGAGGTGAGACAGAGAAACAGCCGCTTGAAGATG 2987
QY 4006 AAGAACCTTACGGAGAGCCAGTCTGTGACACTTGAACATGTCCCAAGCTCAAGAGGGGAGAG 4065
Db 2988 AAGAACCTTACGGAGAGCCAGTCTGTGACACTTGAACATGTCCCAAGCTCAAGAGGGGAGAG 3047
QY 4066 CACTGGGAAGCCTTGGAGAGAGCCCTTCTCTCCAGAGCAAGAGCAAAAGCAGGTTGACTAG 4125
Db 3048 CACTGGGAAGCCTTGGAGAGAGCCCTTCTCTCCAGAGCAAGAGCAAAAGCAGGTTGACTAG 3107
QY 4126 AGGTTCAAGTTCAAGCCTGTGACACAAACAGTCACTCAACAGCAGAGAGCTGTGAAAAG 4185
Db 3108 AGGTTCAAGTTCAAGCCTGTGACACAAACAGTCACTCAACAGCAGAGAGCTGTGAAAAG 3167
QY 4186 TCATAGAAACGGTGTGTGATTTCAAGAGACAGGTGAAGTCCAGAGTGTGTAGTGCACACT 4245
Db 3168 TCATAGAAACGGTGTGTGATTTCAAGAGACAGGTGAAGTCCAGAGTGTGTAGTGCACACT 3227
QY 4246 TATTACCAAGCTGAAGAGTCCCTGTGCAACGGGTGGCCACTGGAACCTTCAGAGTCAAGAG 4305
Db 3228 TATTACCAAGCTGAAGAGTCCCTGTGCAACGGGTGGCCACTGGAACCTTCAGAGTCAAGAG 3287
QY 4306 ACAGGTTACCCCTGGGGGCTGTGAGTCTCAGAGCAGATCCATCAATGATATGTAATCTCTG 4365
Db 3288 ACAGGTTACCCCTGGGGGCTGTGAGTCTCAGAGCAGATCCATCAATGATATGTAATCTCTG 3347
QY 4366 CTCTGAAAGACCTTACATCTTGAACCTTACAAAGAGAAATAGGCAATCCCAAGAGAGAGC 4425
Db 3348 CTCTGAAAGACCTTACATCTTGAACCTTACAAAGAGAAATAGGCAATCCCAAGAGAGAGC 3407
QY 4426 GATCAGAGAGAGAGACAGGCAAGTGTGCTGTGATGTCTGAGGCAAGAGAGGTACAG 4485
Db 3408 GATCAGAGAGAGAGACAGGCAAGTGTGCTGTGATGTCTGAGGCAAGAGAGGTACAG 3467
QY 4486 CAATCGAAAAAGTCTCTCAAGCTGAACCTGAGTCCCTGGAACCTTGAAGGTGAAGAGACA 4545
Db 3468 CAATCGAAAAAGTCTCTCAAGCTGAACCTGAGTCCCTGGAACCTTGAAGGTGAAGAGACA 3527
QY 4546 AGATTGTGCTGAACGTCAATTCAGACAGCCGTTGAACAGTTTGCACGTACAGAAAACAGCCC 4605
Db 3528 AGATTGTGCTGAACGTCAATTCAGACAGCCGTTGAACAGTTTGCACGTACAGAAAACAGCCC 3587
QY 4606 CCGAAACTCATGCTTATGATTTCAAGACCCAGGTTCTGTGATGCAAGGCTTGAACAGAGGG 4665

ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
US-08-978-277A-1

Query Match 76.3%; Score 4701.4; DB 13; Length 5074;
Best Local Similarity 96.8%; Pred. No. 0;
Matches 4988; Conservative 0; Mismatches 76; Indels 91; Gaps 15;

1016 GGAAGAGACAGAGCCAGCTCGAGAGACAGAGCCGCGACAGAGACAGAGCCAG 1075
Db 1 GGAAGAGACAGAGCCAGCTCGAGAGACAGAGCCGCGACAGAGACAGAGCCAG 60
QY 1076 GTTGTACAGACAGCTACAGAGAGTGTGCTCTTTGGAAGACAGGTTGTGACTTGA 1135
Db 61 GTTGTACAGACAGCTACAGAGAGTGTGCTCTTTGGAAGACAGGTTGTGACTTGA 120
QY 1136 GGCATCTCAGAGAGAGAGTGTGCTCTTTGGAAGACAGGTTGTGACTTGA 1195
Db 121 GGCATCTCAGAGAGAGAGTGTGCTCTTTGGAAGACAGGTTGTGACTTGA 180
QY 1196 AGCCACCAAGAGAGTGTGCTCAGAGAGTCCAAGTGAAGACAGAGAGAGAGAGAG 1255
Db 181 AGCCACCAAGAGAGTGTGCTCAGAGAGTCCAAGTGAAGACAGAGAGAGAGAGAG 240
QY 1256 GCAAGGAGAGAGAGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1315
Db 241 GCAAGGAGAGAGAGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
QY 1316 GCGCCCTCAG 1375
Db 301 GCGCCCTCAG 360
QY 1376 GCGATGAG 1435
Db 361 GCGATGAG 420
QY 1436 AAGTCTGAG 1495
Db 421 AAGTCTGAG 480
QY 1496 GGTGTCTCTCAG 1555
Db 481 GGTGTCTCTCAG 540
QY 1556 CTGAG 1615
Db 541 CTGAG 600
QY 1616 CGAAG 1675
Db 601 CGAAG 660
QY 1676 GAGGAG 1735
Db 661 GAGGAG 720
QY 1736 AGGAG 1795
Db 721 AGGAG 775
QY 1796 GAAAG 1855
Db 776 GAAAG 835
QY 1856 AGGAG 1915
Db 836 AGGAG 894
QY 1916 CACCTTGTCTCAG 1975
Db 895 CACCTTGTCTCAG 954

QY 1976 TGAGAGACAAAGCCAG 2035
Db 955 TGAGAGACAAAGCCAG 1014
QY 2036 ACTGATTTGTGTGAGATCATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2095
Db 1015 ACTGATTTGTGTGAGATCATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1071
QY 2096 AGAGAGCCAG 2154
Db 1072 AGAGAGCCAG 1131
QY 2155 AAGAGAGCCAG 2214
Db 1132 AAGAGAGCCAG 1190
QY 2215 GTTCTCAG 2274
Db 1191 GTTCTCAG 1250
QY 2275 CATTAAAGAGATGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2334
Db 1251 CATTAAAGAGATGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1310
QY 2335 AAGAGCTAGATGAG 2386
Db 1311 CCGAG 1370
QY 2387 ATCTTGGGTTTCCATTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2446
Db 1371 ATCTTGGGTTTCCATTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1430
QY 2447 -GCAAG 2505
Db 1431 GCAAG 1490
QY 2506 ATGTCCAG 2555
Db 1491 ATGTCCAG 1550
QY 2556 CCGAG 2625
Db 1551 CCGAG 1607
QY 2626 GTAAG 2685
Db 1608 GTAAG 1667
QY 2686 GTGTGAG 2745
Db 1668 GTGTGAG 1727
QY 2746 CAG 2805
Db 1728 CAG 1787
QY 2806 AAATCTCTGTCTCAG 2865
Db 1788 AAATCTCTGTCTCAG 1847
QY 2866 CCAAG 2925
Db 1848 CCAAG 1907
QY 2926 GATCTGAG 2985
Db 1908 GATCTGAG 1967
QY 2986 TTTCCAG 3045
Db 1968 TTTCCAG 2027
QY 3046 AAGTGTGTGTGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3105

Db 2028 AGGTGGTGTCTAGATACAGAGAGAGGAGCCAGACCGAGGCCATCTCTCAAGCCG 2087
Qy 3106 TTGCAGACAAGGTGAAGAGAGAGTCCAGAGTGCCTGCAACCCAGACTGTGCAGAGAACGG 3165
Db 2088 TTGCAGACAAGGTGAAGAGAGAGTCCAGAGTGCCTGCAACCCAGACTGTGCAGAGAACGG 2147
Qy 3166 GGTCAAAAGCACTGGAGAGGTTGAGAGGTGAGAGGAGGACTCCGAAGTGTGCTTCGG 3225
Db 2148 GGTCAAAAGCACTGGAGAGGTTGAGAGGTGAGAGGAGGACTCCGAAGTGTGCTTCGG 2207
Qy 3226 AGAAAGAGAGGAGCTTTATGCCGAAGAGACCCGTGCAGGAGGCTGGAGCTGAGCATCTTG 3285
Db 2208 AGAAAGAGAGGAGCTTTATGCCGAAGAGACCCGTGCAGGAGGCTGGAGCTGAGCATCTTG 2267
Qy 3286 CACAGGCTCTGAGACTGGACAGGCTACTCCAGAGAGCCTTGAAGTTCTCTGAAGTCAAGG 3345
Db 2268 CACAGGCTCTGAGACTGGACAGGCTACTCCAGAGAGCCTTGAAGTTCTCTGAAGTCAAGG 2327
Qy 3346 CAGATGTAGACCACTGTGCCACAGTGCAGGTTATCAAGCTCCAGCAGCTGATGAAACAGG 3405
Db 2328 CAGATGTAGACCACTGTGCCACAGTGCAGGTTATCAAGCTCCAGCAGCTGATGAAACAGG 2387
Qy 3406 CCGTGGCCCTGAGTCACTCCGAACCTTTGACAGACAGTGAGACAAATGGAGCACTCCCT 3465
Db 2388 CCGTGGCCCTGAGTCACTCCGAACCTTTGACAGACAGTGAGACAAATGGAGCACTCCCT 2447
Qy 3466 TAGCAGATTGACACACTGCGAGATGGGACACAGCAAGATGAACCCATTGACAGCCAGGACA 3525
Db 2448 TAGCAGATTGACACACTGCGAGATGGGACACAGCAAGATGAACCCATTGACAGCCAGGACA 2507
Qy 3526 GTAAAGCCACTGACAGTGTGAGGAGTCAAGGTGACAGAGAAAGAGGGGCTACTGCTC 3585
Db 2508 GTAAAGCCACTGACAGTGTGAGGAGTCAAGGTGACAGAGAAAGAGGGGCTACTGCTC 2567
Qy 3586 AGAAAGAGGAGCCTTCACACTACTTAATATGTTTCAGGCCAGGAGAAACATGGGGAAG 3645
Db 2568 AGAAAGAGGAGCCTTCACACTACTTAATATGTTTCAGGCCAGGAGAAACATGGGGAAG 2627
Qy 3646 AACCAGGAAGAGATGTTCTTTGAACCTACACAGCAAGAGCTTACTGCTGCAGCCGTGCCCG 3705
Db 2628 AACCAGGAAGAGATGTTCTTTGAACCTACACAGCAAGAGCTTACTGCTGCAGCCGTGCCCG 2687
Qy 3706 TTCTGGCAAGACATGAGGTGGGTCAAGAGGTGAGGTGACTGTTGGATGGAGAAAAG 3765
Db 2688 TCTGGCAAAAGACTGAGGTGGGTCAAGAGGTGAGGTGACTGTTGGATGGAGAAAAG 2747
Qy 3766 TCAAAGAGACAGGAGGTGTTGTACACTCTGACCCCAAGTCAAAAGGCTGCTGATG 3825
Db 2748 TCAAAGAGAACAGGAGGTGTTGTACACTCTGAGCCCAAGTCAAAAGGCTGCTGATG 2807
Qy 3826 TGACATATGACAGTGAAGTGAATGGAGTGGCCGGGTGTGAGGAAAAGGAGAGTACTGAAG 3885
Db 2808 TGACATATGACAGTGAAGTGAATGGAGTGGCCGGGTGTGAGGAAAAGGAGAGTACTGAAG 2867
Qy 3886 TGACAGTCTTAGCTCGAGAGGAGAGATGGAACCTGACTGTTGAAAGGAGAAAAGGG 3945
Db 2868 TGACAGTCTTAGCTCGAGAGGAGAGATGGAACCTGACTGTTGAAAGGAGAAAAGGG 2927
Qy 3946 AGACAAAGCCAGACAGTGAAGAGGTGAGCAGGAAACAGCCGCTCTGAGCATG 4005
Db 2928 AGACAAAGCCAGACAGTGAAGAGGTGAGCAGGAAACAGCCGCTCTGAGCATG 2987
Qy 4006 AAGGAACCTACGGGAAGCCAGTCTGTGACATTTGACATGCCAGCTCAGAGGGGGGAGG 4065
Db 2988 AAGGAACCTACGGGAAGCCAGTCTGTGACATTTGACATGCCAGCTCAGAGGGGGGAGG 3047
Qy 4066 CACTGGGAAGCCTTGGAGGAGCCCTTCTCTCCAGACCAAGAACAAGCAGGTTGCAATG 4125
Db 3048 CACTGGGAAGCCTTGGAGGAGCCCTTCTCTCCAGACCAAGAACAAGCAGGTTGCAATG 3107
Qy 4126 AGGTTCAAGTTCAAAGCCTTGACACACAGTCACTCAAAACAGACAGAGCTGTGAAAAGG 4185

Db 3108 AGGTTCAAGTTCAAAGCCTGGACACAAACAGTCACTCAAAACAGCAGAAAGCTGTGAAAAGG 3167
Qy 4186 TCATAGAAAACGGTTGTGATTTTCAGAGACAGGTGAAAAGTCCAGAGTGTGTAGGTGCACACT 4245
Db 3168 TCATAGAAAACGGTTGTGATTTTCAGAGACAGGTGAAAAGTCCAGAGTGTGTAGGTGCACACT 3227
Qy 4246 TATTACCAGCTGAGAAAGTCTCTGCAACGGGTGCCACTCGACTCTTCAGCATGCGAGG 4305
Db 3228 TATTACCAGCTGAGAAAGTCTCTGCAACGGGTGCCACTCGACTCTTCAGCATGCGAGG 3287
Qy 4306 ACACGGTACCCTCTGGGCCCTGAGTCTCAGGCAGAAATCCATCCCAATCATAGTAACTCCTG 4365
Db 3288 ACACGGTACCCTCTGGGCCCTGAGTCTCAGGCAGAAATCCATCCCAATCATAGTAACTCCTG 3347
Qy 4366 CTCTGAAAGCACCTCATCTGACCTACAAGGAGAAATPAAGCGATCCACAGAGAGC 4425
Db 3348 CTCTGAAAGCACCTCATCTGACCTACAAGGAGAAATPAAGCGATCCACAGAGAGC 3407
Qy 4426 GATCAGAGGAGAGGACAAAGCCAGATGCTGCTGATGCTGACGGCAAGGAGATACAG 4485
Db 3408 GATCAGAGGAGAGGACAAAGCCAGATGCTGCTGATGCTGACGGCAAGGAGATACAG 3467
Qy 4486 CAATCGAAAAGTCTCAAGGCTGAACCTGAGATCTCTGGAACTTTGAGAGTAAAGCAACA 4545
Db 3468 CAATCGAAAAGTCTCAAGGCTGAACCTGAGATCTCTGGAACTTTGAGAGTAAAGCAACA 3527
Qy 4546 AGATTGTGCTGAACGTCATTTACAGACAGCCGTTGACCAAGTTTCGCACGTACAGAAACAGCCC 4605
Db 3528 AGATTGTGCTGAACGTCATTTACAGACAGCCGTTGACCAAGTTTCGCACGTACAGAAACAGCCC 3587
Qy 4606 CCGAAACTCATGCTTATGATTTACAGACCAGCTTCTGCAATGCGCTTGGACAGCAGGG 4665
Db 3588 CCGAAACTCATGCTTATGATTTACAGACCAGCTTCTGCAATGCGCTTGGACAGCAGGG 3647
Qy 4666 AGCCACACAGATGCTGACAAATGAAGATGCCAGATGCCAGATGAAACACCCAGTCCGCGAGC 4725
Db 3648 AGCCACACAGATGCTGACAAATGAAGTTCGCAAGATGAAACACCCAGTCCGCGAGC 3707
Qy 4726 CCAGAGAGGACTTGCAGTCTGACCCGTTCTGGAGGATGGGCTCAGCCTCGGAAATGCT 4785
Db 3708 CCAGAGAGGACTTGCAGTCTGACCCGTTCTGGAGGATGG--CTCAGCTCGGAAATGCT 3765
Qy 4786 TGCCGGCTTGCAGTTGAAAAGCCCGGTGTCAAAGTAAAGCATTTGAGAAAGTGTCTCTCA 4845
Db 3766 TGCCGGCTTGCAGTTGAAAAGCCCGGTGTCAAAGTAAAGCATTTGAGAAAGTGTCTCTCA 3825
Qy 4846 ACCCAAAGATCCAAAAGGAGCATGCTGATGGCCCTCAGCTCCAAAGCTTACCCAGG 4905
Db 3826 ACCCAAAGAT-CAAAAAGGAGCATGCTGATGGCCCTCAGCTCCAAAGCTTACCCAGG 3884
Qy 4906 CAGAGGCCAGTGTCTCTGGAACCTTAACCAAAGAAATCCCCAGACACCAACGGACCAAAGC 4965
Db 3885 CAGAGGC--AGTGTCTGGAACCTTAACCAAAGAAATCCCCAGACACCAACGGACCAAAGC 3941
Qy 4966 TAAACCGAGGAGGCGATCCCCCAAAAGTTGAGTCCAGGAAGAAATGCTCTACCAAGT 5025
Db 3942 TAAACCGAGGAGCG-ATGCCCCCAAAAGTTGAGTCCAGGAAGAAATGCTCTACCAAGT 4000
Qy 5026 CAGTCAAAGAAACAAGGCCCAAGCAGAGAGGACCTGCGAGGAGCCAAAGGGAGACTGG 5085
Db 4001 CAGTCAAAGAAACAAGGCCCAAGCAGAGAGGACCTGCGAGGAGCCAAAGGGAGACTGG 4060
Qy 5086 CAGAAATCTTAAGATGTTAGTTGCTATTGTACATCTGTAAAGACAGAAATGTGAAAACAAG 5145
Db 4061 CAGAAATCTTAAGATGTTAGTTGCTATTGTACATCTGTAAAGACAGAAATGTGAAAACAAG 4120
Qy 5146 TCACAGAAACAAGATGCTGCTGTTGGGACCTTGAGACCAAGATTTTCAGAGCCCATGAGATC 5205
Db 4121 TCACAGAAACAAGATGCTGCTGTTGGGACCTTG-GACCAAGATTTTCAGAGCCCATGAGATC 4179
Qy 5206 CAGAGAGCAGGCGCTCCAAATGATTTCCACCCAGATGAGCACCCCGACAAATTCCTGAGGCT 5265
Db 4180 CAGAGAGCAGGCGCTCCAAATGATTTCCACCCAGATGAGCACCCCGACAAATTCCTGAGGCT 4239


```

QY 5266 TCATCGGAGCTAAGAGCCAGCTACATTTCTCTGTTGCAAGACTGCTTTGATTTGCCCC 5325
Db 4240 TCATCGGAGCTAAGAGCCAG----- 4259
QY 5326 TTGATGCGCGTCCGTATTTCTAACATTTCCGTTTCAGACTGCTTTGATTTGCCCC 5385
Db 4260 -----CTAACATTTCCGTTTCAGACTGCTTTGATTTGCCCC 4299
QY 5386 TTGATGCGCGTCCGTATTTCTAACATTTCCGTTTCAGACTGCTTTGATTTGCCCC 5445
Db 4300 TTGATGCGCGTCCGTATTTCTAACATTTCCGTTTCAGACTGCTTTGATTTGCCCC 4359
QY 5446 GCCATACCTAGTTCCACTTCTCAAACTGAGAGATCTCTTATGATTTATATGATG 5505
Db 4360 GCCATACCTAGTTCCACTTCTCAAACTGAGAGATCTCTTATGATTTATATGATG 4419
QY 5506 TTATATGATGCTCTCTCTCTGATACCATTTGATTTCTTAAAGTTAAGCATGCT 5565
Db 4420 TTATATGATGCTCTCTCTCTGATACCATTTGATTTCTTAAAGTTAAGCATGCT 4479
QY 5566 TTATATGATGCTCTCTCTCTGATACCATTTGATTTCTTAAAGTTAAGCATGCT 5625
Db 4480 TTATATGATGCTCTCTCTCTGATACCATTTGATTTCTTAAAGTTAAGCATGCT 4539
QY 5626 GCCTCAACTGATACCTGAGCAAGACATACATTTCTGCGCAAGAACAGATCTTTT 5685
Db 4540 GCCTCAACTGATACCTGAGCAAGACATACATTTCTGCGCAAGAACAGATCTTTT 4599
QY 5686 TTAAGTTTACGATGCTTATGATCTGATGCGCTTATGCTCTGAAAGTGTTGTTCC 5745
Db 4600 TTAAGTTTACGATGCTTATGATCTGATGCGCTTATGCTCTGAAAGTGTTGTTCC 4659
QY 5746 TATGACAGCGAGCTCAGAAATATAAACCCTTTTGAACATCCAGATGTCCTCAAT 5805
Db 4660 TATGACAGCGAGCTCAGAAATATAAACCCTTTTGAACATCCAGATGTCCTCAAT 4719
QY 5806 TACATGATTTTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5865
Db 4720 TACATGATTTTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4779
QY 5866 GTGTCAGATTAAGCCCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5925
Db 4780 GTGTCAGATTAAGCCCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4839
QY 5926 GATGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5985
Db 4840 GATGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4899
QY 5986 GGACCGAATTCATATGAGATGAGAGATGCTCTGTTCTTACAGATGATTTGATG 6045
Db 4900 GGACCGAATTCATATGAGATGAGAGATGCTCTGTTCTTACAGATGATTTGATG 4959
QY 6046 ATACTGAGTTTGTCTGTCTATCTGTGCGCCCTCTTAAAGAACAAATGTCATG 6105
Db 4960 ATACTGAGTTTGTCTGTCTATCTGTGCGCCCTCTTAAAGAACAAATGTCATG 5019
QY 6106 TTCTTTGATTAATTTGATTTGACCACTGATTTAAATTAACATATTTGACTAC 6160
Db 5020 TTCTTTGATTAATTTGATTTGACCACTGATTTAAATTAACATATTTGACTAC 5074

```

```

; LENGTH: 6199
; TYPE: DNA
; ORGANISM: HUMAN
US-60-360-207-12210
Query Match 73.8%; Score 4548.6; DB 80; Length 6199;
Best Local Similarity 86.6%; Pred. No. 0;
Matches 5375; Conservative 0; Mismatches 644; Indels 190; Gaps 25;

QY 4 GCCGCGGAGATGAGAGCCATGAGCCATGAGCGAGCGAGCTTCCAGCAGCGGA 63
Db 113 GCCGCGGAGATGAGAGCCATGAGCCATGAGCGAGCGAGCTTCCAGCAGCGGA 172
QY 64 GCCCGAGCAGCCCGGGAGCGACACGCCGAGCGAGCTGTGTCTCATGTGCCATGGCC 123
Db 173 GCCCGAGCAGCCCGGGAGCGACACGCCGAGCGAGCTGTGTCTCATGTGCCATGGCC 232
QY 124 CCGGAGCTGAAGCTCGGAGCAGCTGAGACCCGCGCAGCGGAGCCCGCCACCAAGC 183
Db 233 CCGGAGCTGAAGCTCGGAGCAGCTGAGACCCGCGCAGCGGAGCCCGCCACCAAGC 292
QY 184 TCCCAAGAGAAATGCGCAGCTGTCTTCTGCAACGCGGTAGCTGAACAGAGATGTC 243
Db 293 TCCCAAGAGAAATGCGCAGCTGTCTTCTGCAACGCGGTAGCTGAACAGAGATGTC 352
QY 244 ATGTTCAGAGAGAAACCAAGAGGGGCGAGAGAGAAAGTGTGTTGATGAGATGTTGAC 303
Db 353 AGTTCAGAGAGAAAGCCAGAGATGGGCAAGAGAAAGTCACTGTGAAAGATGTTGAC 412
QY 304 AGCGAGATCGAAGATGTGAGAGAAAGACGAGTTGAAGAAATGGCGGCAACTCCA 363
Db 413 AGCGAGATCGAAGATGTGAGAGAAAGACGAGTTGAAGAAATGGCGGCAACTCCA 472
QY 364 CAGCTGTGAAGATATCAAAAGATGGGCGAGAGACATCAAGAAATATTGAACAGA 423
Db 473 CAGCTGTGAAGATATCAAAAGATGGGCGAGAGACATCAAGAAATATTGAACAGA 532
QY 424 TCCCTGCTTCAGAAACATGTGAGAAATGTACACCTGCTGAGTCCAGGCTAATG 483
Db 533 TCCCTGCTTCAGAAACATGTGAGAAATGTACACCTGCTGAGTCCAGGCTAATG 592
QY 484 ATGTGCTTCAAGAAATATTTAAATTTGTTGTTTAAATTCACGGTGAAGAGATA 543
Db 593 ATGTGCTTCAAGAAATATTTAAATTTGTTGTTTAAATTCACGGTGAAGAGATA 652
QY 544 AAAATGAAGATCAAGATCTCTCACTACTCTCTCAAGAGATGAAGGCGAAGGG 603
Db 653 AAAATGAAGATCAAGATCTCTCACTACTCTCTCAAGAGATGAAGGCGAAGGG 712
QY 604 CAGAGCTCTGTGTGAGCTGAGACCAACAGAGCCAGTGTGAGACTGCGTGGAG 663
Db 713 CAGAGCTCTGTGTGAGCTGAGACCAACAGAGCCAGTGTGAGACTGCGTGGAG 769
QY 664 AGTACGATCCAAAGAAATGAGCTGAAGCAATCCACAGAGAAAGAGGACCCCTGA 723
Db 770 AGTACGATCCAAAGAAATGAGCTGAAGCAATCCACAGAGAAAGAGGACCCCTGA 829
QY 724 AGCAAGAACAGAGAGCAGCAAGAAATCCCTTCAAGCGAATCTGATCAACGCGTGGAG 783
Db 830 AGCAAGAACAGAGAGCAGCAAGAAATCCCTTCAAGCGAATCTGATCAACGCGGAG 889
QY 784 AAGAAAGCAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 843
Db 890 AAGAAAGCAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 949
QY 844 CCGGAGAGCGCCAGTCAAGTGAAGCAATCTCTTAAAGATTTCTTCACTCAG 903
Db 950 CCGGAGAGCGCCAGTCAAGTGAAGCAATCTCTTAAAGATTTCTTCACTCAG 1009
QY 904 GTTGGGCGGCTGGGCGAAGAGACAGCTTCAAGAAATCAAAAGAGATGATCTGAAA 963
Db 1010 GTTGGGCGGCTGGGCGAAGAGACAGCTTCAAGAAATCAAAAGAGATGATCTGAAA 1069

```

```

RESULT 11
US-60-360-207-12210
; Sequence 12210, Application US/60360207
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig
; TITLE OF INVENTION: HUMAN GENOME DISCOVERY SYSTEM AND USES THEREOF
; FILE REFERENCE: CLO01321
; CURRENT APPLICATION NUMBER: US/60/360,207
; CURRENT FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 47235
; SEQ ID NO 12210

```


Db 3218 CAGACAAAGTAAAGAGACCTCCAGGTGCTGCAACCCAGACTCTGCAAGAGAGGAGG 3277
 Qy 3169 CAAAAGCACTGAGAGAGTTGAGAGGTAGAGAGAGACTCCGAATGTGCTTGAGAGA 3228
 Db 3278 CGAAGAGCACTGAGAGAGGTGAGAGGTAGAGAGAGACTCCGAGGTGCTGAGTACCGAGA 3337
 Qy 3229 AAGAGAAAGAGCTTATGCGCAAAAGACCCGTCAGAGAAAGCTGAGACTGACATCTTTGAC 3288
 Db 3338 AAGAGAAAGATGTTGTGCGGAGAGACCGTGCAGAGAAAGCTGAAACTGAGCATCTTGAC 3397
 Qy 3289 AGGCTCTGAGACTGAGCAGGCTACTCCAGAGAGCCTTGAAGTTCTTGAAGTCAAGGACG 3348
 Db 3398 AGGCTCCGAGACTGTACAGGTTACCCAGAGAGCCTTGAAGTTCTTGAAGTCAAGAGG 3457
 Qy 3349 ATGTAGACCATGTGCGCACGTGCGAGTTATCAAGCTTCAGACGCTGATGAAACAGGCCG 3408
 Db 3458 ATGTAGACCGTGCCACACATGCCAGTTATCAACGACAGCAGCTGATGAAACAGGCTG 3517
 Qy 3409 TGGCCCCCTGAGTCACTCCGAAACCTTGAAGACAGAGTGAACAAATGGAAGCACTCCCTAG 3468
 Db 3518 TGGCCCCCTGAGTCACTGAAACCTTGAAGACAGAGTGAACAAATGGAAGTACTCCCTCG 3577
 Qy 3469 CAGATTGCACTGCTGAGTGGGACACAGCAAGATGAACCTTGAAGCCAGACAGTA 3528
 Db 3578 CAGATTGCACTGCTGAAACCGGACACAGCAAGACGAGACCTTGAAGCCAGACAGTA 3637
 Qy 3529 AAGCCACTGCACTGTCAAGGACGTACAGGTACAGAAAGAGGCGGCTACTGCTCAGA 3588
 Db 3638 ATGCCATTGCGCGCGTCAAGCAGTCAAGGTCACTGAAGAGAGGCAAGTGTGCTCAGA 3697
 Qy 3589 AAGAGAGACCTTCGACACTAATAATATGTTCCAGCCCGAGAAAGAACTGAGGAGAAAGAC 3648
 Db 3698 CGAAGAGGCGCTTCAACACCATC---TAGTTTTCAGCCCGAGAAAGAAACACAGGAAAAAC 3754
 Qy 3649 CAGGAAGAGATGTTGTTGAACCTACACAGCAGAGCTTACTGCTGAGCGGCGCGTTC 3708
 Db 3755 CAGGAAGAGATGTTCTAGAACCCACA---CAAGCGTGTGCTCCGAGGACAGTGTCTATTC 3811
 Qy 3709 TGGCAAGACTGAGGTGGGTCAAGAGGTGAGGTTGACTGTTGAGTGGAGAAAAAGTCA 3768
 Db 3812 TGGCAAAAGGCTGAGGTGGGTCAAGAGGTGAGGCTGGCAGTTGATGAGAAAAAGTCA 3871
 Qy 3769 AAG-----AAGAACAGAGAGTGTGTTGATCACTCTGACCCCACTCAAA 3813
 Db 3872 AAGACGCA CAGTGTGTTAAAGAACTGAGAGTGCCTGTGACACTGAACTCAAA 3931
 Qy 3814 AGGCTGCTGATGTGACATATGACAGTGAAGTGAAGTGGCCGAGGTGTGAGAAAAAG 3873
 Db 3932 AGACTGCTGACTTGACAGTGCAGTGAAGTAAAGAAAGTGGCCAGATGTCAAGAAAACTG 3991
 Qy 3874 AGACTTGAAGTGCAGATCTTACCTGAGAGAGAGAGATGAAACTGACGTTGAAA 3933
 Db 3992 AGAGTAAAGAAAGACAGAGTAAAGTAAAGCCCGAGAAAGAAAGATGGAACCGAGCTTGA 4051
 Qy 3934 AAGAGAAAAAGGAGACAAAGCCAGAGCAAGTGAAGTGAAGAGTGAAGAGAAACAGCCG 3993
 Db 4052 AAGAGAAAAAGAGACCAAGACAGAGCAAGCCAGTGAAGAAACATAGCAGGAAAAAGCTG 4111
 Qy 3994 CTCCGAGATGAAGAAACCTTACGGAAGCCAGTCTGACATTTGACATGAGCCAGTCAAG 4053
 Db 4112 CTCCGAGATGAAGAAACCTTACGGAAGCCAGTCTGACATTTGACATGAGCCAGTCAAG 4171
 Qy 4054 AAGAGGGGAGGACCTGAGAAACCTTGAAGAGAGCCCTTCTCCAGACCAAGCAAAAG 4113
 Db 4172 AAGAGGGGAGGACCTGAGAAACCTTGAAGAGAGCCCTTCTCTCCAGACCAAGCAAAAG 4231
 Qy 4114 CAGGTTGATAGAGGTTCAGTTCAAAAGCTGAGACACACAGTCACTCAACAGCAAGAG 4173
 Db 4232 CAGATTGATAGAGGTTCAGTTCAAAAGCTGAGACACACAGTCACTCAACCAAGCAAG 4291
 Qy 4174 CTGTGAAAAAGGTCAATAGAAAGGTGTTGATTTGAGAGCAGGTGAAGTCCAGAGTGG 4233
 Db 4292 CTGTGAAAAAGGTCAATAGAAAGGTGCAACTTCAAGAGTGAAGAAAGTTTGGAGTGTG 4351

Qy 4234 TAGTGCACACTTATTACAGCTGAGAAAGTCTTGCAGAGGCTGGCCACTGCACTCTTC 4293
 Db 4352 CAGGTGGGCATATTCACAGCTGAGAAAGCTCTCGAAACCGGTGGCTACGGGACTCTTC 4411
 Qy 4294 AGCATGCGAGAGACAGGTTACCCCTGGGGCGTGAAGTCTCAGGCGAGATTCATCCCAATCA 4353
 Db 4412 AGCATGCGAGAGACAGGTTACCCCTGGGGCGTGAAGTCTCAGGCGAGATTCATCCCAATCA 4471
 Qy 4354 TAGTAACTCTGCTCTGAAAAGCAACCTTACATCTGACCTAACAGAGAAATTAAGCGAT 4413
 Db 4472 TAGTAACTCTGCTCTGAAAAGCAATCTTACATCTGACCTTCAAGAGAAAGTGGCGCAT 4531
 Qy 4414 CCGAGAGAGACCATTCAGAGAAAGAGACCAAGCAAGTGTGCTGCTGATGCTACCGCA 4473
 Db 4532 CCGAGAAACAGAGATCAGATGAAGATTAACAGCCAGATGTGCTGATGCTGATGCTCGCA 4591
 Qy 4474 AGGAGATGACAGCAATGGAATAAGTCTCAAGAGCTGAACCTGAGATCCTGGAACTTGAGA 4533
 Db 4592 AGGAGATGACAGCAAGAGAGAAATCTCTAGAGGCTGAACCTGAGATCTTGAACTTGAGA 4651
 Qy 4534 GTAAGAGCAACAAATGTGTGCTGAACGTCAATTGACAGCCGTTGAACAGTTGACAGTA 4593
 Db 4652 GTAAGAGCAATTAAGATTTGTCAGAGTGTCAATCCAGACGCGTGACCAAGTTGACAGTA 4711
 Qy 4594 CAGAAACAGCCCCGAAACTCATGCTTATGATTCAGACAGCCAGTTCTG-CATGCAAG 4652
 Db 4712 CAGAAACAGCCCCGAAACCCACGCTTCTGATTTACAGAAATCAGATTTCTGTGATGACAG 4771
 Qy 4653 CTGACAGCAGGAGGCCCAACAGATGTGTGACAAATAATGAAG-----AT 4697
 Db 4772 CTGACAGCAGGAGGACACAAAGATGTGACAAAGATGAAGAGGACCTTCAAGTCTGCC 4831
 Qy 4698 GCCAAGTGAATAACCCAGATGCGCAGCGCCAGAGAGACTTGCAGAGTCTGACCGTTCTG 4757
 Db 4832 CCGAAGATGAACACTCAGTGCCTGACAGAGAGGACTTGGGTTTCTGATGTTCTG 4891
 Qy 4758 GAGCATGGGCTCAGCTCGGAAATGCTTGCCTGCGCTTGCAGTTGAAGCGCGGTGCA 4817
 Db 4892 AAGCATGTGACAAAGCTTCAAAATGATCAACGCTTGCAGTTGAAGTGCCTGATGCA 4951
 Qy 4818 AAGTAAAGATTAAGAAAGTGTCTCTCAACCCAAAGATCCAAAGAGAGCATGCTGAT 4877
 Db 4952 AAGAAAGTGTGAAGAGTGTCTCTTCAAGTCAAGAT-GAAAAGAAACATGCTGTGAC 5010
 Qy 4878 GGCCCTCAGCTTCAAAAGTTAGCCAGGACAGGCGCAGTGTCTGGAACCTTAACAA 4937
 Db 5011 GGCCCCGACGACCAAAAGTTAGCCAGGACAGGCGGATGCTCTGGAAATCTTAACAA 5070
 Qy 4938 GAATCCCAACACCAACGACCAAAAGTAAACGAGAGGCGCATCCCCCAAAAGTTGAG 4997
 Db 5071 GAGTCCCAACACCAACGACCAAAAGTAAACGAGAGGAGATGCTCTGGAAGAAAGAA 5130
 Qy 4998 GTCCAGAGAAAGAAATGTCTACCAAGTCACTCAAAAGAAACAAGGCCAGAGCAAGAGAG 5057
 Db 5131 AT-----GAACAAAGGCCAGACAGAGAG 5154
 Qy 5058 ---GACTGACAGAGCCAAAGGAGACCTGAGCAATCTTAAGATGTTAGTGTCTATTG 5114
 Db 5155 GACGACCTACAGAGGACCAAGGAGACCTGACAGATCTTAAGACGTGAAGTGTGCTCAT- 5212
 Qy 5115 TACATCTGTAAGCCAGATGTGAATAACAGTCAACAGAACAAAGTGTGCTGTTGGAGCC 5174
 Db 5213 -----TGTAAATCTAGAAATGTAAGTGAACAGAACAAAGTGTGCTGTTGGAGCC 5266
 Qy 5175 TTGAGACCAATTTTCAAGAGCCCATGAGATCCAGAGAGCAGGCGCTTCAATGATTTCCA 5234
 Db 5267 TTGAGACCAAAATTTTCAAGAGCCTTGAAGTGCAGAGACCAAGCGCTCAATGATTTCAA 5336
 Qy 5235 CCCAGTAGACACCCGCAATTTCTGAGGCTTCAATCGGAGCTGAGAGCAGCTAACTATTT 5294
 Db 5327 CCTGAGAGACCCCGCAATCTGAGGCTGCAATCGGAGCTGAGAGCAGCTAACTATTT 5386

Db	1035	ACTAGTCCCGTGACCAAGTGAACAGGATCAACCTTCAAAAAATCTTACTCAAGGTTGG	1094
Qy	909	GGCGGCTGGGCGCAAGAAAGACACAGCTTCAAGAAATCAAAAAGAGATGATCTGGAAACTGCC	968
Db	1095	GCCGGCTGGGCGCAAAAAGACCAAGTTCACAGAAAGCCGAGAGAGATGAAGTGGAAAGCTTCA	1154
Qy	969	GAGAAAGAAAAGAGCAGAGAGCAGAAAAAGTAGACAGAGAAAGAAAAGAAAAGACAGAG	1028
Db	1155	GAGAAAGAAAAGAAACAAGAGCCAGAAAAAGTAGACAGAAAGAAAGAGGAAAGGCAGAG	1214
Qy	1029	CCAGCTTCGAGGA-----CGAGAGCCGGCAGGAAGAC	1061
Db	1215	GTTCCTCCGAGAAACTGACCGGCTCCGAGCAAGCCACCCACAGAGACCCGCGAGAAAGT	1274
Qy	1062	ACAGACCAGGCAAGGTTGTGACAGACACTACGAGAAAGTGGAGCTGCTTTGGAAAGACAG	1121
Db	1275	GCCCAACGAGCCCGGTTATATACGCTGAATATGAAAAGTTGAGCTCCTCAGAGAGACGA	1334
Qy	1122	GTTCGTGACCTTGAGGAGCAGTCGTACAGAGAAATGTGCTCCTTGGCAAGAAAGTGT	1181
Db	1335	GTCAAGTGGCTGCAGGGGACCTTCTGAAGAGAAACCTGCTCGTTGGCAACAGAAAGTGT	1394
Qy	1182	GATGAGAAATGGAAGCCCAAGAA---GTTCGTGACAGAGTCCACGTGACACCTGTG	1238
Db	1395	GATGAGAAAATAGAAAGTCCACCAAGAAAGAGTTGTGGCCGAAAGTCCACGTACAGACCGTG	1454
Qy	1239	GAGAAAGCAGAGGAGGAGCAGGGGAGGAGGAGAGAGAGTGAAGGGGGGCGTGGTGGTAGA	1298
Db	1455	GAGGAGAGAACCGAAGAGCAAAA-----ACGAGAGTGA	1490
Qy	1299	GAAGCAGAGAGATCTTGTGCCCCCTGAGAAACTGGCTGAGCCCGCAGAGGTCCTCCAGAA	1358
Db	1491	GAAGCAGAGAGGTCGTGTGCCAGCTGAAGAAATGTGGTAATGATGACGAACCTCAAGAA	1550
Qy	1359	GCTGAGCTGCTGAGAGCTGATGAAAGACAGAGAAATGTGTGTCTCTGAGAGAGACAC	1418
Db	1551	GCCGAAGCTGCGCAAGGAGCTGTGTGAAGCTCAAGAAACGTGTGTTCCGAGAGAGACCT	1610
Qy	1419	ACTCAACTGACAGACCTTAAGCTCGAAGAAAGAACGCTCCCAAAACCCCAAGAGGACTT	1478
Db	1611	ACACAGGAGGTGACTCTCAAGCTCTGATGABAAGTGTCTTCCAAACCCCGAGAGGCTT	1670
Qy	1479	GTCAGTGAAGTGAGAGTGTCTCTCTCAGAGAAAGATCAAGTACAGGAGAGTCCCTTG	1538
Db	1671	GTGAGTGAAGTGAAGATGTCTGTATCAACAGAGAAATGAAGTGAAGGAAAGTCACTA	1730
Qy	1539	AAGAAACTTTCAGTACCTCAGGCTTAAAGAGCTGTCTGGGAGAGACGAGGGGAAA	1598
Db	1731	AAGAAAGCTTTTACACACACTGCGCTTAAAAAACCTTCTGGAAGAAAGAACAGAAAGGAAA	1790
Qy	1599	CGAGAGGTTGGGGAGACGAAAGAGCCTGAGAAATACCAACACATTCACACGGAATCCCCA	1658
Db	1791	AGAGGA---GAGAGAGACAGAGAAATCAGGGAGACACACTCAGGTTCAAGCGAATTTCCG	1847
Qy	1659	GAGAGTGTGTATGAGCAAGAGGAGAGAGCTGTGCGTGTGTCCTCCGAGAGAGCTGAGAG	1718
Db	1848	GACAGCCAGAGAGAGCAAAAAGGGGAGAGCTGTGCTCATCCCTGAGGAGCCGAGAG	1907
Qy	1719	ACCACGTGTCTTGAGAAAGGCGCTGGAAGCACCCCGAGATGGGGAGCTGAGAGGA	1778
Db	1908	ATCAGGTGTCTGAAAAAGGGCTTAGCCGAGAGTCAACAGAGATGGGGAGCTGACMAAGGA	1967
Qy	1779	ACTACTTCCGATGAGAGAAAGAAAGAGAGGATCACTCCCGGGCAATCTTCAAAAAG	1838
Db	1968	GCTACTTCCGATGAGAGAAAAAAAAGGAAGGTGTACTCTCCGGGCATCACTTCAAAAAG	2027
Qy	1839	ATGAGTCAACCCCAAGAAACGGGTCCGAAGACTTTCAGAGGTGACAGAGAGAGAGACTG	1898
Db	2028	ATGAGTCAACCCCAAGAGCGTGTTAGAGCGGCTTCGAAAGTGTATTAAGAAATGAGCTG	2087
Qy	1899	GAGAAAGTCAAGAGCGCAACCTGTCTCCACTGATAGCACAAGTGCAGAAATGCAGAT	1958

Db	2088	GACAAAGTC	CAAGACCGCT	ACTTCCTTC	CTCCACCGAGAC	ACAGCCTCTG	AAATGCAAGA	2147		
Qy	1959	GAAGTC	AAAACTGTTG	TGAGAA	CAAAAGCC	AGAGAA	CCAAAGCCTG	AGGTGATACT	2018	
Db	2148	GAAT	TGAAAGGAG	CGTGGA	AGAGCC	AAAGCCGGA	AAAGAACCGCAAG	TGATGATACC	2207	
Qy	2019	TCAGT	CTTGGGA	GCATG	ATTGTGT	CGGAT	TCATCCAA	GAAGAGCAAGGAAGCA	2078	
Db	2208	TCAGAT	CTTGGGA	AGCTTTAA	TTTGTGTG	GGATAT	CATCCAGAAA	AAAGCAAGGAAGAGG	2267	
Qy	2079	TCCT	CTTCAGAT	GTGAAG	AGGGCC	CAAGAC	ACTGGAG	GGGAGCACTGCACAGAG	2138	
Db	2268	TCCT	CTTCTGAT	TGAGAA	GGGGG	AGCCAAA	GCAT	TGGAGAGAC	CACAGAAAGCTGAT	2327
Qy	2139	GAGG	CAGCAAA	AGA	CAAAAGAG	CCGAG	CAGACG	CTGTTCTCTCCAG	CACCAGAGACAG	2198
Db	2328	GAGG	CCGGA	AAAAAG	CAAAAGAG	CGGGG	CAGACG	GGGATCTTGCTGTGTTCC	CAAGAACAT	2387
Qy	2199	GACCA	AGGCGCA	AGGAAGT	TCTTCA	CCCCGAG	CCAGCGGG	AGACCCCTTCCGA	AGGGGAAGT	2258
Db	2388	GATC	CAGGGCGAG	GGGAAT	CTCTCC	CGAGCAAG	CTGGAAG	CCCTTACGGA	AGGGGAGGCG	2447
Qy	2259	GTCT	CTCACTTGG	AGTCA	TTTAAAGT	TATG	CACTCC	CAAGAAAAATCC	CAAGTCAAA	2318
Db	2448	GTTT	CTCCTG	GGAGT	CA	TTTAAAG	GT	TATGTCAG	CCAGAAAAATCAA	2507
Qy	2319	CTGGA	AGAAAG	CCGGAAG	AC-----	TC	TAGT	TGAGCA	GTGTCCACTGAG	2366
Db	2508	CTGGA	AGAAAG	CAAAAG	CAAGCT	CCAT	TAGCTGG	GTGGTGTG	AGCAATTCACATCC	2567
Qy	2367	ATCGA	ACCGAGT	ATAGAA	AGAA	TTCTTGG	TTTCCAT	TAAAGAA	TTTATCTCCCGAG	2426
Db	2568	ACTGA	ACCCGGT	AAAGAA	AGAT	CTTGG	GTCTCA	TCTCA	AAAGATTTTATCTCTG	2627
Qy	2427	AAGAA	AGGCGAG	ACGGGAG	CAAGAA	CAAGCC	ACTGTGA	AGACTC	AGGSCAGTGGAG	2486
Db	2628	AAGAA	AGGCGAG	ACGGGAG	CAAGAA	CAAGCC	CTGTG	AAAGAC	CGAGGGCC	2687
Qy	2487	AT	AAATGAG	AG	CGA	CCCTA	TATG	TCCAG	CCGTCGTCTGTG	2546
Db	2688	GC	CAACGA	AAATG	AT	CTG	ATG	TCTCCG	CCGTCGTCTCTGTG	2747
Qy	2547	GAG	AGGAGAA	GATG	-----	AA	CC	CA	GGGAA	2537
Db	2748	GAA	AGGAGAA	ATG	AGG	AC	CAG	CAAGCC	CAAAAAAG	2807
Qy	2598	GGG	GTGTG	TAC	GTGTG	CCGAG	AGCTC	ATG	AGAGCTG	2657
Db	2808	GC	AGC	CACTG	AGGTG	TC	CA	AGAG	CTC	2867
Qy	2658	GT	CAT	TGA	TGGG	AC	AGG	AGT	CAC	2717
Db	2868	GT	CG	TGA	CGG	AG	GC	AT	CA	2927
Qy	2718	GCTT	CCG	TAC	GA	A	CTT	TGA	AC	2777
Db	2928	GCTT	CA	TG	GA	CA	CA	CTT	TGA	2987
Qy	2778	GT	C	A	T	G	A	A	A	2834
Db	2988	GA	T	T	G	A	A	A	A	3047
Qy	2835	GAGG	T	A	A	A	A	A	A	2894
Db	3048	GAG	A	C	A	G	A	G	G	3107
Qy	2895	GT	G	A	C	A	G	C	T	2954
Db	3108	GT	G	A	C	A	G	C	T	3167
Qy	2955	GCC	G	A	A	A	A	A	A	3014
Db	3168	GCT	G	A	A	A	A	A	A	3227

QY 3015 ACAGAGAACCCACCCAGTTTCAGGAGGTAGAGTGGTGTCTAGATACAGAGAAGAG 3074
DB 3228 ACAGAGAGCCACTCCCGTGCAGGAGGTGGAAGGTGGCGTACCTGACATAGAAGACAA 3287
QY 3075 GAGCGCCAGACGACGAGGCATCTCCAAAGCCGTTTCAGACAAAGGTGAAAGAGGAGTCCAG 3134
DB 3288 GAGGCGGAGTCAAGAGTCTCCAGGAGTCCAGGAGTGCAGAAAAGTGAAGAGGAATCCAG 3347
QY 3135 GTGCTGCAACC---CAGACTGTGCAGAGAACCGGGTCAAAAGCACTGAGAAAGTTGAG 3191
DB 3348 CTGCTGACACCGGTGGGCGCAGAAAGTGTGCTTCAGCCTGTGCAGAGAGCAGAGGCAGAA 3407
QY 3192 GAGGTAGAGAGCACTCCGAAGTCTGGCTTCGAGAAAGAGAGAGCTTATCCGAAA 3251
DB 3408 AGACCAAGAGAGCGGCTGAAGCGTGGGCTGAAGAAAGAGACGGATGTAGTGTGAAA 3467
QY 3252 GGAACCGCTGCAGGAAGCTGCAGCTGAGCATCTTCAGACAGGCTCTGAGACTGGACAGGCT 3311
DB 3468 GTAGATGCTCAGGAGGCAAAACCTGAGCCTTTTACACAGGGAAGGTGGTGGGCGACCC 3227
QY 3312 ACTCCAGAGACCTTG---AAGTCTCTGAAGTCAAGGAGATGTAGAC-----CAT 3359
DB 3528 ACCCAGAAAGCTTTGAAAAGCTCTCTCAAGTCAAGAGAGCATAGAGTCCAGTGAGCTT 3587
QY 3360 GTCCGACAGTGCCA-----GGTTATCAAGCTCCAGCACTGATGGAA 3401
DB 3588 GTAACCACTTGTCAAGCCGAAACCTTAGCTGGGGTAAATCAAGAGAGATGGTGTATGGAA 3647
QY 3402 CAGGCGGTGGCCCTGAGTCATCCGAACCTTCAGACAGTGCAGACAAATGGAAGCACT 3461
DB 3648 CAGGCTATCCCCCTGACTCGGTGGAACCCCTACAGACAGTGCAGTGTAGTGAAGCACC 3707
QY 3462 CCCTTAGCAGATTCAGACACTGCAGATGGGACACAGCAAGATGAACCATTTGACAGCCAG 3521
DB 3708 CCGTAGCCGACTTTGACGCACAGGCACAAACCCAGAAAGACAGATTTGGAAATCCAT 3767
QY 3522 GACAGTAAAGCCACTGAGCTGTGAGGAGTCAAGGTTCAGAGAACAGAGCGGCTACT 3581
DB 3768 GAGGAGAATGAGGTGCGCATCTGGTACCCAGTCAAGGGGCAAGAGCAGAGCGAGTTCTCT 3827
QY 3582 GCTCAGAAAGAGGCTTCGACACTACCTAATAATGTTCCAGCCAGGAGAACATGGG 3641
DB 3828 GCACAGAAAGAGGCTTCAGACACTTCAGACTTTGTGTTGTCAGGAAGAACTAAAGAA 3887
QY 3642 GAAGAACCCAGGAAGAGATGTTTGAACCTTACAGCAAGAGCTTACTGTGCGAGCCGTG 3701
DB 3888 CAATCAAAGATGGAAGACACTCTAGAGCATACAGATAAAGAGGTGTCAGTGGAAACTGTA 3947
QY 3702 CCGTGTCTGCAAAAGACTGAGGTGGGTCAAGAGGGTGA-----GGTTGACTGGTTG 3752
DB 3948 TCCATTCTGTCAAAAGACTGAGGGGACTCAAGAGGCTGACCAGTATGCTGATGAGAAACC 4007
QY 3753 GATGGAGAAAGACTCAAGAGAACAGGAGGTGTTGTACACTCTGG-----ACCCAAC 3806
DB 4008 AAAGACGTACCACTTTTCGAAGGACTTGAAGGGTCTATAGACACAGGCATPAACAGTCAGT 4067
QY 3807 AGTCAAAAGGCTGCTGATGTGACATATGACAGTGAAGTGTGGAGTGGCGGGTGTGAC 3866
DB 4068 CGGGAAGAGTCACTGAGTTGGCTTTAAGGTGAAGGGACAGAGAGCTGAATGTAA 4127
QY 3867 GAAAAGAGAGTACTGAAGTGCAGAG-----TCTTAGCCTGGAGGAGGA 3911
DB 4128 AAGGATGATGCTCTTGAATGCAGAGTCAAGCTAAGTCTCTCCATCCCGGTGGAGAGA 4187
QY 3912 GAGATGGAAGTCACTGTTGAAAAGGAGAAAGGAGACAAAGCCAGAGCAAGTGAAGTAA 3971
DB 4188 GAGATGGTGTCAAGTTCGAAGAGGAGAAACAGAAAGCAGAGCAACCCATGTGAATGAA 4247
QY 3972 GAAGG---TGAGCAGGAACACAGCGCTCTCTGAGCATGAAGAACCTACGGGAAGCCAGTC 4028
DB 4248 GAGAACTTGAGCACGAAACAGCTGTTTACCGTATCTGAAGAGGTGAGTAAAGCAGCTCCTC 4307

QY 4029 CTGACACTTGAACATGCCAGCTCAGAGAGGGGGAAGGCACCTGGGAAGCCCTTGGAGGAAGC 4088
DB 4308 CAGACAGTGAATGTGCCCATCATAGATGGGCAAGAAAGTACAGCAGTTTGGAAAGAAC 4367
QY 4089 CTTTCTC---TCCAGACCAAGACAAAGCAGGTTGCATAGAGGTTCAAGTTTCAAGCCTG 4145
DB 4368 CTTCTTCCCTCCCTAGGTCAAGAGGAGGAGATGACCAAAATTCAGTTTCAGAGCTCT 4427
QY 4146 GACACAACAGTCACTCAAAACAGCAGAAAGCTGTGAAAA-----GGTCAATAGAAAAGGTT 4199
DB 4428 GAGGCATCATTCATCTAAACAGCGGTGCAGAGGAGGAAAAGGCTTTCAGGAGAACTGCC 4487
QY 4200 GTGATTTTCAGAGACAGGTGAAAGTCCAGAGTGTGTAGGTGCACACTTATTACAGCTGAG 4259
DB 4488 AACATTTTGAAGAACAGGTGAAACGTTTGGAGCCTTGCAGGTGCACATTTAGTTTTCGGAAGAG 4547
QY 4260 AAGTCTCTTCGAACGGGTGGCCACTCGACTCTTTCAGCATGTCAGAGGACAGGTTACCCCTG 4319
DB 4548 AATCTCTTGAAAAAATGAAGACTTTTGGCTCATTCAGGGGAAGATGCTGTGTCCCAACA 4507
QY 4320 GGGCTTGAAGTCTCAGGCAGAAATCCATCCCAATCATAGTAATCTCTCTCTCTCTGAAAGCACC 4379
DB 4608 GGGCCCGACTGTTCAGGCAAAATCGACACAGTGATGTATCTGCTACTACCAAGAAAGGC 4667
QY 4380 CTATCTCTGACCTCAAGAGAGAAATAAGGCAATCCAGAGAGCGATCAGAGGAAGAG 4439
DB 4668 TTAAGTTCGGAACCTGGAAGGAGAGAAAACCATCACTGAAGTGAAGTCAGATGAAGTC 4727
QY 4440 GACAAGCCAGATGCTGCTGATGCTGACGCGCAGGAGAGTACAGCAATCGAAAAGTC 4499
DB 4728 GATGACAGGTTGCTTGGCC-----AGGAGGTCAAAAGTGAAGTGTAGCAATTTGAGGAGGAT 4781
QY 4500 CTCAAGGCTGAACCTCAGATCTCTGGAATCTTGAGAGTAAAGAGCAACAAAGATTGTGCTGAAC 4559
DB 4782 TTAGACCTGAAAATGGGATTTGGAACTTGAGACCAAAAGCAGTAAACTTGTGCCAAAC 4841
QY 4560 GTCATTTACAGACAGCCCTTGACCAAGTTTCGACGTPAC---AGAAAACAGCCCCCGAAACTCAT 4616
DB 4842 ATCATCCAGACAGCCCTTGACCAAGTTTGTACGTTACAGAAAGAAACACCCGAAATGTTG 4901
QY 4617 GCTTATGATTCACAGACCCAGGTTG-----CTGCAATCAGGCTTCACAG 4660
DB 4902 AGCTCTGAGTTACAGACACAAAGCTCAGTGTATAAAGCTGCAGAGCCAGCGCTGGACAG 4961
QY 4661 CAGGAGCCCAACAGATGCTGGAACAAAATGAAAGATGCAAGATGAAACACCCAGTGCC 4720
DB 4962 GAAAACGGAAGAGAGGAGGAGGAACTCAGGCTCTGCAAGATGAAACACCAATTAAT 5021
QY 4721 GCAGCCCA-----GAGAGGACTTGCAGAGTCTGACCGTCTCTGGA 4759
DB 5022 TCAGCCAAAGAGAGGAGTCAGAGTCAACCCGAGTGGGACAAAGCATTTCTGATATTTCCAAA 5081
QY 4760 GGCATGGGCTCAGCCTCGGAAATGCTTGGCGCTTGCAGTTGAAAGCGCGGTGTCAA- 4818
DB 5082 GACATGAGTGAAGCCTCAGAAAAGCAATGACTGTTGAGGTAGAAGGTTTCCACTGTAAAT 5141
QY 4819 ----- 4818
DB 5142 GATCAGCAGCTGGAAGAGTGTCTCCCATCTGAGGAAGAGGAGGTGGAGCTGGAACA 5201
QY 4819 -----AGTAAGCATTTGAGAGAGTGCCT 4840
DB 5202 AAGTCTGTGCCAGAGATGATGTGTCATGCTTTGTAGCAGAAAGATAGAGAGTCACTA 5261
QY 4841 CTTCAACCCAAAGA---TCAAAAGAGAGCATGCTGTGATGGCCCTCAGCTCCAAAGCTTA 4898
DB 5262 GTTGAACCCAAAGAGATGAAAAGGTGATGATGTTGATGACCTGAAAACCCAGAACTCA 5321
QY 4899 GCGCAGGAGAGGCGCAGTCTCGAATCTAAACCAAGAAATCCCCAGACACACACGGA 4958
DB 5322 GCCTTGGCTGATCTGATGCTCAGAGGCTTAAACCAAGAGGTCCCCAGATACAAATGGA 5381
QY 4959 CCAAAGCTTAACGAGGAGGCGGATCCCCCAAAAGTTGAGGTCCAGGAAGAAAGATGTCT 5018

Db 583 GTTCACGACATCAGATGATGGCAGGAGGAGACCCCGAAATAATCGAACATTCCT 642
Qy 429 GCTTCAGAAACAATGTGGAAGAAATGGTACAGCCTGCTGAGTCCCGAGCTAAATGATGTT 488
Db 643 TCTTCAGAAAGCAATTTAGAGAGCTAAACAACCCACTGAGTCCCGAGCTAAATGATATT 702
Qy 489 GGCTTCAAGAAAGTATTTAAATTTGTTGGTTTAAATTTCAAGTGAAGAGGATAAAAT 548
Db 703 GGAATTAAGAAAGGTGTTAAGTTTGTGGCTTTAAATTTCACTGTGAAGAAAGGATAAGACA 762
Qy 549 GAAAGTCAAGATCTGTCCAACTACTCACTGTCAAGAGGATGAAGGCGAAGGGCGAGAA 608
Db 763 GAGAGCCTGACACTGTCCAGCTACTCACTGTGAAGAAAGTGAAGGGGAGGAGCAGCA 822
Qy 609 GCCTCTGTGAGGTGTGAGACCAACAGAGCCAGTGTGAGAGATGCGGTGCGAGAGTCA 668
Db 823 -----GGGGGTGGCGACCAACAGACCCCGAGCCTTGGGGCT-----GGAGAAGCA 867
Qy 669 GCATCCAAAGAAAGTGAAGCAATCCACAGAGAGCAAGAGGACCCCTGAAGCAA 728
Db 868 GCATCCAAAGAAAGCGAAACCCAAACAATCTACAGAGAAACCCGAGAGACCCCTGAAGCGT 927
Qy 729 GAAACAGACGACACAGAAATCCCTCTCAAGCCGAATCTGTATCAAGCGGCTGAGGAAGAA 788
Db 928 GAGCAAGCCAGCGAGAAATTTCTCCCGACCGAATCTGGCCAAAGCA--GTGGAGGAA 984
Qy 789 GCCAAAGANTGAAGAGAGAAAGAAACAAGAGAAAGAGCCCAAGTCCCGAGAAATCCCG 848
Db 985 TGCAAGAGGAGAGAGAGAGAAACAAGAAAGAAACCTAGCAAGTCTGCAGAAATCTCCG 1044
Qy 849 AGCAGCCAGTCAACAGTGAACACATCTTCTTCAAGAGTCTTCACTCAGGTGG 908
Db 1045 ACTAGTCCCGTGAACAGTGAACAGGATCAACCTTCAAAATAATTTCTCACTCAAGGTGG 1104
Qy 909 GCCGGCTGGCCAAAGAGACAGCTTCAAGAAATCAAAAGAGGATGATCTCGAAACTGCC 968
Db 1105 GCCGGCTGGCCAAAGAGACAGCTTTCAGAGCGGAGAGGATGAAGTGGAAAGCTTCA 1164
Qy 969 GAGAAGAAAGGAGCAGAGCGAGAAAGAAAGTAGACAGGAGAAAGAAAGAACACAGAG 1028
Db 1165 GAGAAGAAAGGAAACAAGAGCCAGAAAGTAGACACAGAAAGAGCGGAAAGCGCAGAG 1224
Qy 1029 CCAGCTCGGAGGA-----GGAGGAGCGCGGAGAGAC 1061
Db 1225 GTTCCCTCCGAGAAATCTGACCGCTCCGAGCAAGCCCAACACAGAGCGCGCAGAAAGT 1284
Qy 1062 ACACACAGGCGCAGTGTGACAGACATACAGAGAGGTGAGCTGCCCTTTCGAAAGACAG 1121
Db 1285 GCCACAGGCCCCGTTATCAGCTGAATATGAGAAAGTTGAGCTGCCCTCAGAGGAGCAA 1344
Qy 1122 GTTGGTGACCTGGAGGATCGTCAGAGAGAGTGTGCTCTCTTTGGCAAACGGAAGTGT 1181
Db 1345 GTCAAGTGGCTCGCAGGACCTTCTGAAGAGAAACCTGCTCCGTTGGCGACAGAGTGT 1404
Qy 1182 GATGAGAGATGGAAGCCACCAAGAA--GTTGTTGAGAGGTCACGTCAGACCCGTCG 1238
Db 1405 GATGAGAAATAGAAAGTCCACCAAGAGAGGTTGTGGCCGAAGTCCACGTCAGCACCGTG 1464
Qy 1239 GAGAAGACAGAGGAGGAGCAGGAGGAGGAGAGGCTGAAGGGGGGCTGCTGCTAGAA 1298
Db 1465 GAGGAGAGACCGGAAGAGCAGAA-----ACGAGGTGGA 1500
Qy 1299 GGAAACAGGAGAAATCTTGGCCCTCGAGAACTGGCTGAGCCCGCAGGAGGTCCCCAGGAA 1358
Db 1501 GAAACAGCAGGGTCTGTGCCAGCTGAAGATTGGTTGAATGGATGCAGAACCTCAGGAA 1560
Qy 1359 GCTGAGCCTGTGAGGAGCTGATGAAGACAGAGAGATGTGTCTCTGAGGAGAGACAC 1418
Db 1561 GCCGAAACCTGCAAGGAGCTGGTGAAGCTCAAGAAACGTCGTGTTTCCGAGAGGAGCCCT 1620
Qy 1419 ACTCACTGACAGACCTTAAGTCTGAGAGAGAGCTGCCCAACACCCAGAGGCAATT 1478
Db 1621 ACAAGGAGCTGACCTCAGTCTGATGAGAAAGGTGCTGTCNAAACCCCGAAGGCGTT 1680

Qy 1479 GTCAGTGAGGTGAGATGCTGTCTCTCAGGAAAGAAATCAAGGTAACAGGAAGTCCCTTG 1538
Db 1681 GTGAGTGGAGTGGAAATGCTGTCTATCAGCAGAGAGATGAAGGTGCAGGAAGTCCACTA 1740
Qy 1539 AAGAAACTCTTCTAGTACTCAGGCTTAAAGAGCTGTCTCGGAAGAGCAGAGGGGAAA 1598
Db 1741 AAGAGCTTTTACCGACTGGCTTAAAGAGCTTCTGGAAGAAACAGAAAGGAAA 1800
Qy 1599 CGAGGAGGTGGGGAGACGAGAGCCTGGAGAAATACAAACACATTTCAACCGAATCCCCA 1658
Db 1801 AGAGG---GGAGGAGACGAGGAATCAGGGGAGCACACTCAGGTTCCAGCCGATTTCTCCG 1857
Qy 1659 GAGAGTCTGATGAGCAAGAGGAGAGAGCTCTGCGTGTCTCCCGCAGAGACCTTGAGAG 1718
Db 1858 GACAGCCAGGAGGAGCAAAAGGGCGAGAGCTCTGCTCTCATCTCCCTGAGGAGCCCGAGAG 1917
Qy 1719 ACCACGTGTCTGAGAAAGGGCGCTGTAAGACACCCAGAGATGGGAAGCTGAGGAAGGA 1778
Db 1918 ATCAGTGTCTGGAAGAGGGCTTAGCCGAGGTGAGCAGGATGGGGAAGCTGAAGAAGGA 1977
Qy 1779 ACTACTTCCGATGGAGAGAAAGAGAGAGAGGATCACTCCCTGGGCACTCTTCAAAAAG 1838
Db 1978 GCTACTTCCGATGGAGAGAAAGAGAGAGTGTCACTCCCTGGGCTTCTTCAAAAAG 2037
Qy 1839 ATGCTGACACCAAGAAACGGGTCCGAAAGACCTTCTGAGAGTGACAAAGAGGAAGAGCTG 1898
Db 2038 ATGCTGACGCCCAAGAGCGTGTGTAGACGGCTTTCGGAAGTGTATAAGAAAGATGAGCTG 2097
Qy 1899 GAGAGGTCAAGAGCCACTTGTCTCTCACTGATAGCAGTGTCCAGAAATGCAAGAT 1958
Db 2098 GACAAGTCAAGAGCGCTACTTGTCTTCCACCGAGAGCAGCCTCTGAAATCAAGAA 2157
Qy 1959 GAAGTCAAAACTGTGTGTGAGGAAACAAAGCCAGAGAAACAAAGCGTAGGGTGGTACT 2018
Db 2158 GAATGAAGGGAGCGTGTGAAGCCAAAGCCGGAAGAACCAAGCCAGAGTGGATACC 2217
Qy 2019 TCAGTGTCTTGGAGACACTGATTTTGTGTGCGATCATCCAAAGAGAGAGCAAGGAAGCA 2078
Db 2218 TCAGTATCTTGGGAAGCTTTAAATTTGTGTGGATCATCCAAAGAAAGAGCAAGAGAGGG 2277
Qy 2079 TCCTCTCAGATGATGAAGAGGGCCNAGGACACTGGGAGGGNACAGTCAAGAGCAGAG 2138
Db 2278 TCCTCTTCTGATGAGGAGGGGACCAAAAGCAATGGGAGGAGACCAACCAAGAAAGTGT 2337
Qy 2139 GAGGCCAGCAAGACAAAGAAAGCCGGAACAGAGCTGTCTCTGCCAGCACCCAGAGAGCAG 2198
Db 2338 GAGGCCGAAAAGACAAAGAGACGGGGACAGACGGGATCTTGTCTGTTCCCAAGAACAT 2397
Qy 2199 GACCAAGCCGAAGAAAGTTCCTCAACCGAGCCAGCGGGAAGCCCTTCCGAAGGGGAAGGT 2258
Db 2398 GATCCAGGGCAGGGAAGTTCCTCCCGGAGCAAGCTGGAAGCCCTACCGAAGGGGAGGGC 2457
Qy 2259 GTCTCCACTTGGGAGTCATTTAAAGATTAGTCACTCCAAAGAAAAAATCCAAGTCAAAA 2318
Db 2458 GTTTCACCTGGGAGTCATTTAAAGGTTAGTCAACGCAAGAAAAAATCAAAGTCCAAG 2517
Qy 2319 CTGGAAGAGAAAGCCGAAGAC-----TCTAGTGTAGAGCAGTGTCCACTGAG 2366
Db 2518 CTGGAAGAGAAAGCCGAAGACTCCATAGCTGGGTCTGTGTAGAACATTTCCACTCCAGAC 2577
Qy 2367 ATCGAACCGAGTAGAGAGAAATCTTGGGTTTCCATTAAAGAAATTCATCCCGGACGCGG 2426
Db 2578 ACTGAACCCCGTAAAGAGAAATCTCTGGGTCTCAATCAAGAAAGTTTATTTCTGTGACGAAG 2637
Qy 2427 AAGAAAGGCGACAGCGGAGAGCAAGAACCACTGTGGAAGACTCAGGCGCAGTGGAG 2486
Db 2638 AAGAAAGGCGCAGATGGGAAACAAAGAACAGCCCTGTTGAAAGACGAGGCGCCACAGGG 2697
Qy 2487 ATAAATGAGGACGACCTAATGTCCAGCGCTGTGCTCTGTCTGATGATTAATGCAAGT 2546
Db 2698 GCCAACGAAGATGACTCTGATGTCCCGCGCGTGGTCCCTCTGTCTGATGATGATGCTGTA 2757

QY 2547 GAGAGGAGGAAGATGG-----AAGCCAGGGGAATACGGAGCTGCCACGCTGCTG 2597
 Db 2758 GAAAGGGGAAATATGAGGACACAGAACCCAAAAAGCGCAGAGCAGCCGAGCAAG 2817
 QY 2598 GGGGCTGTGTAGTGTCCGAGAGGCTCAGTAAGACTGTGTCCACTGTGATGCCA 2657
 Db 2818 GCAGCAGCTAGAGTGTCCAGAGAGCTCAGCCGAGACTCAGTTTCAATATGATGACAGAGT 2877
 QY 2658 GTCATTTGATGGAGCAGGGCAGTCAACAGTGTGAAAGCGGTCTCTTGATATCC 2717
 Db 2878 GTCCGTGACGGGACAGGGGAGCTACCATTTATGAAAGAGTCTCTTCTTGATATCT 2937
 QY 2718 GCTTCCGTACAGAACCTTTGAAACACACAGGGGGAAGCCATCCACTGTGAAG 2777
 Db 2938 GCTTACGTACAGAACCTTTGAAACAGTGAAGCTGAAGCCGACCTGTTAACTGAGAG 2997
 QY 2778 GTCACTGAAAAAGACATCATTTG---AGNAGAACTCCGTGTCTCACCCAGAGCTTACCA 2834
 Db 2998 GTATTTGAAAAAGAGATATTTGCAAGAAAGAACCCCCACGGTTACTGAACTCTGCA 3057
 QY 2835 GAGGTTAAGATGCCCATGACGACATGCTCACCAAGTGAAGTGAATTTCACTCAGAGCT 2894
 Db 3058 GAGAACAGAGAGGCCCGGGGCGACACGGTCTTATGAGGCGGAATTGACCCCGAAGCT 3117
 QY 2895 GTGACAGCCACAGACACTCAGAGGCTCTCCGTACTGAAGAAATTACGGAACATCGGG 2954
 Db 3118 GTGACAGCTGCAAGAACTGCAAGGCGCATTTGGGTGCGAAGAAAGAACCGAACATCTGT 3177
 QY 2955 GCGCAAGAGACACACAGACATGCTGCCGAGTTTCCAGCTGACTGACTCCACAGACACC 3014
 Db 3178 GCTGAAGAGACACAGAAATGCTGTACAGAGTCTCCAGTTAACGACTCCACAGACACC 3237
 QY 3015 ACAGAGAGGCCACCCCAAGTTGAGAGGTGTGTGCTAGATACAGAAAGAG 3074
 Db 3238 ACAGAGAGGCCACTCCGGTGAGAGGTGGAAGGTGGGTACTGACATAGAAAGCAA 3297
 QY 3075 GAGGCCACAGCCAGGCACTCTCCAGCCGTTTGACAGACAAAGTGAAGAGAGTCCGAG 3134
 Db 3298 GAGAGGCGGACTCAAGAGGTCTCTCAGACAGTGGCAGAAAAAGTGAAGAGAAATCCAG 3357
 QY 3135 GTGCTGCAACC---CAGACTGTGCAAGAACGGGGTCAAAAGCACTGGAAGAGTTGAG 3191
 Db 3358 CTGCTTGGGACCGGTGGGCGCAAGATGTGCTTCAAGCTGTGCAAGAGCAAGAGCAGAA 3417
 QY 3192 GAGGTAGAGAGGACTCCGAAGTGTGCTTGGAAGAAAGAGAGAGCCTTATGCCGAAA 3251
 Db 3418 AGACAGAGAGGACGCTGAAAGCTGCGGGTCTGAAGAAAGAGACGGATGTATGTTGAAA 3477
 QY 3252 GAGCCGTGCAAGAGCTGAGACTCTTGACACAGGGCTTGAGACTGAGACAGGCT 3311
 Db 3478 GTAGATGCTCAGAGAGGCAAAAACTGAGCTTTTACCAAGGGAAGGTGGTGGGAGAC 3537
 QY 3312 ACTCCAGAGAGGCTTG---AAGTTCCTGAAGTCAAGGCGAGATGTAGAC-----CAT 3359
 Db 3538 ACCCAGAGAGGCTTTGAAAAAGCTCTCAAGTCAAGAGAGCATAGAGTCAAGTGTAGCTT 3597
 QY 3360 GTGCGCACGTGCA-----GGTTATCAAGCTCCAGCAGCTGATGAA 3401
 Db 3598 GTAACCATTTGTCAAGCCGAAACCTTAGTGGGTAAATATCACAGAGATGTGATGAA 3657
 QY 3402 CAGGCGGTGGCCCTTGAGTCTCCGAAACCTTGACAGACATGAGACAAATGGAAGCACT 3461
 Db 3658 CAGGCTATCCCTCTGACTCGGTGAAACCCCTACAGACAGTGAAGCTGATGGAAGCAC 3717
 QY 3462 CCCTTAGCAGATTCAAGACATGAGATGGGACACAGCAAGATGAACAATTCAGCCAG 3521
 Db 3718 CCGGTAGCCGACTTTGACGACACAGGACACACCAAGAAAGAGAGTTGTGAATTCAT 3777
 QY 3522 GACAGTAAAGCCACTGACAGTGTCAAGGAGTCAAGGTCAAGAAAGAGAGCGGCTACT 3581
 Db 3778 GAGGAGATGAGGTGCACTGTGTACCAAGTCAAGGGGGGACAGAAAGCAAGGAGTTCCT 3837
 QY 3582 GCTCAGAAAAGAGAGGCTTTCAGACTAATTAATTTTCAGCCCGAGGAAGAACATGGG 3641

Db 3838 GCACAGAAAGAGAGGCTCCAGACCTTCCAGTTTGTGTTCAGAGAAAGAACTTAAAGAA 3897
 QY 3642 GAAGAACCCAGAGAGATGTTTCTTGAACCTACACAGAGAGCTTACTGTGACCGGTG 3701
 Db 3898 CAATCAAAAGATGAAGACACTTAGAGCATFACAGATPAAAGAGGTGTGATGAAACTGTGA 3957
 QY 3702 CCCGTTCTGGCAAGACTGAGGTGGGTCAAGAGGTGA-----GTTGACTGGTTG 3752
 Db 3958 TCCATTTCTGTAAGACTGAGGGGACTCAAGAGGCTGACAGATGCTGATGAAGAAAAAC 4017
 QY 3753 GATGAGAAAAAAGCTCAAGAAAGACAGAGGTGTTTGTACCTGTG-----ACCCAAC 3806
 Db 4018 AAAGACGTACATTTTGTGAAGACTTGAGGGGTCTATAGACACAGCATPACAGTCAAT 4077
 QY 3807 AGTCAAAAGGCTGTGATGTGACATATGACAGTGAAGTATGGAAGTGGCGGGTGTGAG 3866
 Db 4078 CCGGAAAAAGTCACTGAAGTGGCCCTTAAAGGTGAAGGGAAGAAAGCTGAATGTATAA 4137
 QY 3867 GAAAAAGAGTACTGAAGTCCAGAG-----TCTTAAGCTTGAGAGAGGA 3911
 Db 4138 AAGATGATGCTCTTGAACCTGACAGATCAGCTAAGTCTCTTCATCCCGGTGAGAGA 4197
 QY 3912 GAGATGGAACCTGACGTTGAAAAAGAGAAAAAGGAGCAAAAGCCAGAGCAAGTGAAGAA 3971
 Db 4198 GAGATGTGATTTCAAGTCAAGAGGAGAAAAAGAGAAAGAGCAAGAGCCACCATGTGATGAA 4257
 QY 3972 GAAAG---TGAGCAGAGAAACAGCCGCTCTTGAGCATGAAGGAACCTACGGGAACCACTC 4028
 Db 4258 GAGAAAGCTTGAGCAGAAACAGCTGTTAACGTATCTGAAGAGGTCAATGACAGCTCTC 4317
 QY 4029 CTGACACTTGAACATGCCCAGTCAAGAGGGGGAAGGCACTGGGAAGCTTTGAGGAAGC 4088
 Db 4318 CAGACAGTGAATGTGCCATATATGATGGGCAAGAAAGTCAACAGTTTGAAGGAAGC 4377
 QY 4089 CCTTCTC---TCCAGAGCAAGACAAAGCAGTTGCACTAGAGTTTCAAGTTCAAAAGCTG 4145
 Db 4378 CTTCTCTCTGCTTAGGCTCAAGAGAGGAGGAGTATGACACCAAAATTCAGTTCAAGCTCT 4437
 QY 4146 GACACAAAGTCACTCAACAGCAGAACTGTGAAAA-----GGTCATPAAACGGTT 4199
 Db 4438 GAGGATCATTTCACTTAAACAGCGGCTGCAAGAGGAAAAAGTCTTAGAGAAACTGCC 4497
 QY 4200 GTGATTTCAAGACAGGTGAAGTTCAGAGTGTGTAAGTCACTTATACAGCTGAG 4259
 Db 4498 AACATTTTGAAGAACAGGTGAACGTTGAGGCTCAGGTGCACTTTAGTTCTGGAAGAG 4557
 QY 4260 AAGTCTCTGCAAGGGGTGGCCACTGACCTTTCAGACATGACAGAGACACGGTACCCCTG 4319
 Db 4558 AATTCCTTGAAAAAAATGAAGACTTTGCCGCTCATCCAGGGGAAGATGCTGTGCCACA 4617
 QY 4320 GGGCTGAGTCAAGGCAAGTCAATCCATCATATGTAATCTCTGCTCTGAAAGCAC 4379
 Db 4618 GGGCCGCACTGTCAAGGCAAAATTCGACACAGTGAATGATCTGCTACTTCAAGAAAGC 4677
 QY 4380 CTACATCTGACCTTCAAGAGAAATTAAGCCATTCACAGAGAGCATCAGAGGAAG 4439
 Db 4678 TTAAGTTCGACCTGGAAGGAGAGAAAAACACATCACTGAAGTGAAGTCAAGTGAAGTTC 4737
 QY 4440 GACAAAGCAGATGTGCTGCTGATGCTGACGGCAAGGAGATGACAAATGAAAAAGTTC 4499
 Db 4738 GATGAGCAGGTGCTTGCC-----AGAGGTCAAAAGTGAATGTAGCAATTTGAAGAGAT 4791
 QY 4500 CTCAAGCTGAACCTGAGATCTGGAACCTTAGAGATPAGAGCAACAGATTTGTGCTGAC 4559
 Db 4792 TTAGAGCTTGAAGATGGAATTTTGAATTTGAACTTGAACCAAGAACTGTCCAAAAAC 4851
 QY 4560 GTCATTTCAAGACCGCTTGACCAAGTTCAGACTAC---AGAAACAGCCCGGAACTCAT 4616
 Db 4852 ATCATTCAGACAGCCGTTGACCAAGTTTGTACGTACAGAAAGAAACGACCCGAAATGTTG 4911
 QY 4617 GCTTATGATTCACAGACCCAGTTT-----CTGATCAGAGCTTGAAG 4660

Db 4912 ACGTCTGAGTTACAGACACAAAGCTCACGTGATAAAAGCTGACAGCCAGGAGCGCTGGACAG 4971
Qy 4661 CAGGAGCCCAACAGATGCTGGACAAAAATGAAAGATGCGCAAGATGAAACACCCAGTGCC 4720
Db 4972 GAAACGGAGAAAGAGGAGAGAACTCAGGCCCTCTGCACAGGATGAAACACCAATTA 5031
Qy 4721 GCAGCCCA-----GAGAGACTTGCAGTCTCAAGTCTCTGACCGTCTCGGA 4759
Db 5032 TCAGCCAAAGAGGAGTCAAGTCAACCGCAGTGGGCAAGCACATCTCTGATATTTCACAA 5091
Qy 4760 GGCAATGGGCTCAGCCTCGGAAATGCTTCCCGCTTTCAGTGTGAAAGCGCGGTGTCAA- 4818
Db 5092 GACATGAGTGAAGCCTCAGAAAGACCATGACTGTGTGAGGTAGAAGTTCACATGTAAAT 5151
Qy 4819 ----- 4818
Db 5152 GATCAGCAGCTGGAAGAGGTCGTCTCCATCTCAGGAAGAGGAGGTGGAGCTGGAACA 5211
Qy 4819 -----AGTAAGCATTGAGAACTGGCT 4840
Db 5212 AAGTCTGTGCAGAAAGATGATGGTCATGCCCTTGTACAGAAAGAAATAGAGAAGTCACTA 5271
Qy 4841 CCTCAACCCAAAGA--TCCAAAGAGCATGCTGCTGATGGCCCTCAGCTCCAAAGCTTA 4898
Db 5272 GTTGAAACGGAAGAGATGAAGAGGTGATGATGTGATGACCTTGAAACCCAGAACTCA 5331
Qy 4899 GCCAGCAGAGGCCAGTGCCTCTGGAAACCTAACCAAGAAATCCCGACAGACCAACGGA 4958
Db 5332 GCCCTGGCTGATCTGATGCCCTCAGGAGGCTTAACCAAGAGTCCCGCAGATACAAATGGA 5391
Qy 4959 CCAAGCTAACCGAGGAGGCGATCCCGCAAAAGTTGAGGTCCAGGAGAAAGAAATGTCT 5018
Db 5392 CCAAAACAAAAGAGAGGAGGATGCCAGGAAGTGAATTTGCAGGAGGAAAGTGCAC 5451
Qy 5019 ACCAAGTCAGTCAAGAGAAACAAGGCCAGCAGAGAGGACCTGCAGGAGCCCAAGGGA 5078
Db 5452 AGTGAATCAGATAAGCGATCACCCACAGCAGAGGAGTTACAGAAACAAGAGAGA 5511
Qy 5079 GACCTGCAGAAATCCTAAGATGTTAGTTG-----CTCATTTGTAC 5117
Db 5512 GAATCTGCAAGTCAGAACTTACAGAACTTAAACATCATGCAAGTTAAACTCATTTGCT 5571
Qy 5118 ATCTGTAAAGCCAGAAATGTGAACCAAGTACAGAAACA-----GATGCTGCTGGGAC 5173
Db 5572 GTTTGGAAGACCAAGATGTGAAGACAAGTAGTAGAAGAAATGAATGCTGCTGAGA- 5630
Qy 5174 CTTGAGACCAAGATTTTCAGAGCCCATGAGATCCAGAGAGGAGGCGCTCCCAATCATTTCC 5233
Db 5631 CTGAAGACCAAGATTTTGA-ACCTTGAGAAATGGAGAGCGGCACATCACTGATCTCA 5689
Qy 5234 ACCAGTAGACACCCGACAAATTCAGGCTTCATCGGAGCTTAGAGCCAGCTTAACATT 5293
Db 5690 TTTCTAGAGAGC-CCCTGACAATCCTGAGGCTTCATCAGAGCTAGAGCCATTAACTT 5748
Qy 5294 TCCTCGTTTCAAGACTGCTTTGATTTGGCCCTTGAATGCGCGTCCGCTGATTTTCAACATT 5353
Db 5749 TCCTCTTTTCAAGACCAACCTACAATTTTCCCTTGAAC----- 5788
Qy 5354 TCCTCGTTTCAAGACTGCTTTGATTTGGCCCTTGTATGTCGCTGATTTTCGATTTA 5413
Db 5789 -----CATATAAATTTCTGATTTA 5806
Qy 5414 AGGTCTCGCTTCAACCTGGAAC-CAATTCCTGCCATACCTAGTTCACCTTCTCAAACT 5472
Db 5807 AGGTCTTAATTTCTTAACCTGGAAGTGGAGTTGGCAATACCTAGTTCTGCTTCTGAAC 5866
Qy 5473 GGAGCATCCTCTTATGATATTATGATGATGTTTTATGATGCTCTCTGCTGACCTAT 5532
Db 5867 GGAGTATCAATCTTTACATATTTATGATGATGTTTTAAGTAG---TCTCTCTGATCTAT 5923
Qy 5533 TGTATATTTTTTCTACAGTTT-----AAGCATGCTTTTTGTATTA 5575
Db 5924 TGTATATTTTTTCTTAATGTTAAGGAAATGTGCAGGATACATGCTTTTTTGTATCA 5983

Qy 5576 TGCAATATATAACGGGTGTGCAGCCATAGCAGCGCTTTTGAAGCTCCAAGCCTCAACTG 5635
Db 5984 CACAGTATATGATGGGCAATGTGCCATAGTGCAGGCTTGGGAGCTTTAAGACCTCAGTTA 6043
Qy 5636 TAACCTCGCAGCAACAGATAACTTC-----CTGGCAAGAGAGACAGTCTTTTT 5686
Db 6044 TATAACCCACAAAAACAGAGCCTCTAGATGTAACATCTCTGATCAAGGTACAAATTTCT 6103
Qy 5687 TAAAGTTTACTGATGCTTAGATCTGTGGGCTTCTAGTCTCTGAAAGTGGTTGTTTCTCT 5746
Db 6104 TAAATTTCACTAATGATTTAGGTCCATATTTAGTGGTACTCTGAAATTTGGTCACTTTCT 6163
Qy 5747 ATGCACAGCAGCTCAGAAATAAAAAACCCCATTTTGAACATCCAGATGTCCCAATATT 5806
Db 6164 ATTACACGGAGTGTGCCAAAACTAAAAAGCATTTTGAACATACACAGAAATGTTCTATTGTC 6223
Qy 5807 ACCATGATTTTTTCCCCCTTTTGTCTAATCCAGTCCAGGTGGAAAGAGTCTCTCTG 5866
Db 6224 ATTGGGAAATTTTGC-----TTTCTAACCCAGTGGAGGTAGAAAGAAAGTTATTTCT 6276
Qy 5867 TGTGAGA-----TTAAGCCCTGTCTCTTTAAATGATATGACAAATCAGT 5909
Db 6277 GTTAGCAAAATTAACCTTTTACATCCTTTTCTACTTGTATGTTGTTGGACCGATAAGT 6336
Qy 5910 GTGCTTAAGCCCATGAGA-----TGTTCTTAATGCAGAAAGAAATCTGT 5953
Db 6337 GTGCTTAATCCTGAGGCAAGTAGTAATATGTTTATATGTTATGAAGAAAGAAATGT 6396
Qy 5954 TGTACGTTTTTTTGTATGTTACTCTTATCTGACCGAAATTCATATGCAGATCGAAGTG 6013
Db 6397 TGTAAAG--TTTTTGAATCTACTCTTATATGCTGACTGCAATTCACATGCGCATGAAATA 6454
Qy 6014 AGTCTGTTCTTTACAGATGTTATTTGATAGACTGTTGAGTGTCTGTTATATATCTG 6073
Db 6455 AGTCAGTTCTTTACAAATGTTATTTGATAGACTGTTGTTGTTGTTGTTGTTGTTGTTG 6514
Qy 6074 TGCCCTTCTTTAAGAACAAATGTTGATTAATGTTCTTTTGGATAAAATTTGATTGACAA 6133
Db 6515 TGCCATTTCTTTAAGAACAAATGTTGCAACAACTTCAATTTGATAAGTTGTTGATTGACGA 6574
Qy 6134 CTGATTTAAATAAA 6147
Db 6575 CTGATTTAAATAAA 6588

RESULT 14
US-60-245-228-517/c
; Sequence 517, Application US/60245228
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN CYCLASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN CYCLASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CLO00878
; CURRENT FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 630
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 517
; LENGTH: 6607
; TYPE: DNA
; ORGANISM: HUMAN
US-60-245-228-517

Query Match 35.4%; Score 2178; DB 68; Length 6607;
Best Local Similarity 64.4%; Pred. No. 0;
Matches 4222; Conservative 0; Mismatches 1795; Indels 537; Gaps 43;

Qy 3 GGCGGGGGGAGTAGAGCCACTGAGCCATGGCGCAGCAGTTCACCGGACGCG 62
Db 6443 GGCTAGCGCGGAGAGTAGCGAGGAGCCATGGCGCGGAGCTCCACCGAGCGC 6384

Db	5321	GCCACAGCCCCGGTTATCAGCTGAATTGAGAAAGTTGAGCTGCTCCAGAGAGCA	5262
Qy	1122	GTTCGTGACCTGAGGCATCGTCAAGAGAAAGTGTGCTCTTTGGCAACGGAAGTGT	1181
Db	5261	GTACGTGCTCGAGGGACCTTCTGAAAGAAACCTGCTCGTTGGCGACAGAGTGT	5202
Qy	1182	GATGAGAAATGAGACCCCAACAA--GTTGTTGACAGAGTTCAGTGAACACCTG	1238
Db	5201	GATTAGAAATAGATGTCCACAAAGAAAGGTTTGCGCGAAGTCCAGTCACAGCCGT	5142
Qy	1239	GAGAAACAAGAGAGCAGGAGAGAGAGAGAGGCTGAAGGGGGGCTGTGTGTGA	1298
Db	5141	GAGGAGGAACCGAAGAGCAAAA-----ACGAGGTGAA	5106
Qy	1299	GAAACAGAAATCTTTCGCCCCCTGAGAACTGCTGAGCCCAAGAGTCCCCAGAA	1358
Db	5105	GAAACACAGGCTGTGTGCAGCTGAAGATTGTTGAATGGAATGCAAACTCAGAA	5046
Qy	1359	GCTGAGCTTGCTGAGAGCTGATGAAGACAGAGAGATGTGTCTCTGAGAGACAC	1418
Db	5045	GCCGAACTCCCAAGAGGCTGTGAAGCTCAAAAGAAAGTGTGTTTCCGAGAGACCT	4986
Qy	1419	ACTCAATGACAACTAGTCTCTGAAGAGAGAGCTGCCCCAACCCGAAGGCAAT	1478
Db	4985	ACACAGGAGCTGACCTCAGTCTGTATGAAAGGTGTGTCCAAACCCCGAAGGCT	4926
Qy	1479	GTCAGTAGGTGAGATGTGTCCTCTCAAGAAATCAAGTACAGGAGTCCCTTG	1538
Db	4925	GTGAGTGAAGTGAATGCTGTATCAGAGAGAGATGAAGGTGCAAGAACTCACTA	4866
Qy	1539	AAGAACTCTTCACTGCTCAGGCTTTAAAGAACTGTCTGGAGAGACAGAGGGAA	1588
Db	4865	AAGAGCTTTTACAGCACTGCTTAAAAAGCTTCTGAAAGAAACAGAAAGGAAA	4806
Qy	1599	CGAAGAGTGTGGGGAAGCAAGAGCCTGAGAAATCAACATTTCAACCGAATCCCCA	1658
Db	4805	AGAGAA--GAGAGAGCGAGAAATCAGGGAGACACTCAGGTTCCAGCCGATTTCTCG	4799
Qy	1659	GAGAGTCTGATGAGCAAGAGGAGAGAGCTGTGCGTGTCCCCCGAGAGCTGAGAG	1718
Db	4748	GACAGCAGAGAGGAGCAAAAGGCGAGAGACTGTGCTCATCCCTGAGAGACCCGAGAG	4699
Qy	1719	ACCACTGTCTGAGAAAGGCGCTGGAAGCACCCAAGATGGGAACTGAGAGGA	1778
Db	4688	ATCACGTGTCTGAAAAAGGCTTAGCCGAGGTGCACAGAGATGGGAACTGGAAGGA	4629
Qy	1779	ACTACTTCGATGAGAGAAAGAGAGAGGATCACTCCCTGGGCACTCCTCAAAAG	1838
Db	4628	GCTACTTCCATGAGAAAAAAGAAAGGTGTCACTCTCTGGGCAATTCAAAAAG	4599
Qy	1839	ATGCTGACACCCAAAGAAACGGGTCACAAACCTTCTGAAGTGCACAGAGAAAGACTG	1898
Db	4568	ATGTGAGCCCAAGAAAGCTGTTTAAACGCTTGGAAAGTGAATGAAGATGAGCTG	4509
Qy	1899	GAGAAAGTCAAGAGCCGACCTTGTCTCTCACTGATAGCAAGTGTCAAAATGCAGAT	1958
Db	4508	GACAAAGTCAAGAGCGCTTACTTGTCTTCCACCGAGACACAGCTTGAATGCAANA	4449
Qy	1959	GAACTCAAACTGTGTGTGAGAAACAAAGCCACAGAGAACAAACGTGAGGTGATCT	2018
Db	4448	GAAATGAAGGAGACGTGGAAGGCCAAGAGCCGAAAGCAACAAACGCAAGGTGATACC	4389
Qy	2019	TCAGTGTCTTGGGAAGCACTGATTTGTGTGCTGATCTCCAAAGACAGACAGAAAGCA	2078
Db	4388	TCAGATCTTGGAAAGCTTTAATTTGTGTGGATCATCCAAABAAAGACAGAGAGGG	4329
Qy	2079	TCTCTTCAATGATGAAGAGGGCCAGAGACCTGGAGGGGACAGTCAACAGACAGAG	2138
Db	4328	TCTCTTCTATGAGAAAGGGGACCAAAAGCAATGGAGAGACCAACAGAAAGCTGAT	4268
Qy	2139	GAGGCAAGAAAGCAAGAAAGCCGAGACAGAGCTGTCTTCCAGACCAACGAGAGAG	2198

Db 4268 GAGCGCGGAAAGACAAAGAGACGGGGACAGACGGGATCCTTGTGGTTCCCAAGAACAT 4209
Qy 2199 GACCAAGCGCAAGAAAGTTCTTCACCCAGAGCAGCGGAAGCCCTTCGAAAGGGGAAGGT 2258
Db 4208 GATCCAGGCGCAGGGAAGTTCTCTCCCGAGCAAGCTGGAAGCCCTTACCGAAGGGAGGGC 4149
Qy 2259 GTCTCCACTTCGGAGTCATTTAAAGATTAGTCACTCAAGAAAATAATCAAGTCAAAA 2318
Db 4148 GTTTCACCTGGAGTCATTTAAAGGTTAGTCAAGCCCAAGAAAATAATCAAGTCCAAAG 4089
Qy 2319 CTGGAAGAGAAAGCCGAAGAC-----TCTAGTGTAGACAGTGTGTCACACTGAG 2366
Db 4088 CTGGAAGAGAAAGCGAAGACTCCATAGCTGGGTCTGTGTAGAACATTCACATCCAGAC 4029
Qy 2367 ATCGAAACCGAGTAGAAGAAATCTTTGGGTTTCATTAAAGAAATTCATCCCGGACGGCG 2426
Db 4028 ACTGAACCCGGTAAAGAAAGATCTCTGGGTCTCAATCAAGAAAGTTTATCTCTGGACGAAGG 3959
Qy 2427 AAGAAAAGGCGAGCGGAGCAGCAAGCAAGCCACTGTGGAGACTCAGGSCCAGTGGAG 2486
Db 3968 AAGAAAAGGCCAGATGGGAAACAAGAACCAAGCCCTGTTGAAGACGCGAGGSCCAACAGGG 3909
Qy 2487 ATAAATGAGGACGACCCCTTAATGTGCCAGCGTGTGTCCTCTGTCTGAGTATAATGCACTG 2546
Db 3908 GCCAACGAAGATGACTCTGATGTCCCGCGCTGTGTCCTCTGTCTGAGTATGATGCTGTA 3849
Qy 2547 GAGAGGGAGAAAGATGG-----AAGCCAGGGGAAATACGGAGCTGCCCCAGCTGCTG 2597
Db 3848 GAAAGGGAGAAAATGGAGGACAGCAAGCCCCCAAAAAAGCGCAGAGCAGCCCGAGCAGAA 3789
Qy 2598 GGGCTGTGTACGTGTCCGAGGAGCTAGTAAGACTCTGTGTCACACTGTGAGTGTGCGCA 2657
Db 3788 GCAGCCACTGAGGTGTCCAAAGGAGCTCAGCGAGAGTCAGGTTTCATATGATGGCAGCAGCT 3729
Qy 2658 GTCAATTGATGGACACAGGCGAGTCACCAAGTGTGCAAGACGGTCTCTCTCTGTGATATCC 2717
Db 3728 GTCCGTGACGGGACAGGGGAGCTACCAATTAATGAAGAAAGGTCTCTCTCTGTGATATCT 3669
Qy 2718 GCTTCGTGTACAGAACCTCTTTGAACACACACAGCGGGGAGAACCATGCCACCTGTTGAAGAG 2777
Db 3668 GCTTCAGTGAACAGACCTCTTTGAACAAAGTAGAAGCTGAAGCCGCACTGTTAACTGAGGAG 3609
Qy 2778 GTCACTGAAAAGACATCATTTGC---AGAAGAACTCTGTGCTCACCCAGACGTTTACCA 2834
Db 3608 GTATTGGAAGAGAGTAATTTGCAAGAGAAAGAACCCCCACGGTTACTGAACTCTGCGCA 3549
Qy 2835 GAGGGTAAAGATGCCATGACACATGTCACCAAGTGAAGTGAATTTACCTCAGAAAGCT 2894
Db 3548 GAGAACAGAGGCGCGGGGCGACACGGTCTGTTAGTGGCGGAATTTGACCCCCGAAGCT 3489
Qy 2895 GTGACAGCCACAGAGACCTCAGAGGCTCTCCGTACTGAAGAAAGTTACCGAAGCATCGGG 2954
Db 3488 GTGACAGCTGCAGAAACTGCAGGGCCATTGGGTGCCGAAGGAACCGAAGCATCTGCT 3429
Qy 2955 GCCAAGAGACCAAGACATGTTGTCGAGTTTCCAGAGTGTACCTCCCGAGACACC 3014
Db 3428 GCTCAAGAGACCAAGACAAATGTTGTACAGCTCTCCAGCTTTAACCGACTCCCGAGACACC 3369
Qy 3015 ACAGAGAGACCCACCCAGTTTCAGAGGTAGAGAGTGTGTGTAGATACAGAGAGAGAG 3074
Db 3368 ACAGAGAGGCCATCTCCGGTGCAGAGGTGAAGGTGGCGTACTGTGACATAGAAGACAA 3309
Qy 3075 GAGCGCCAGACGAGGCGCATCTCCAAAGCCGTTTCAGACAAAGGTGAAGAGGAGTCCAG 3134
Db 3308 GAGAGCGGACTCAAGAGGTCTCTCCAGGAGTGCAGAAAAGTGAAGAGGAATCCAG 3249
Qy 3135 GTGCTTGAACC---CAGACTGTGCAGAGAACGGGGTCAAAAGCACTGTGGAAGAGTTGAG 3191
Db 3248 CTGCTTGGCACCGGTGGGCCCAAGAGATGTGCTTCAGCTGTGTCAGAGCAGAGCGCAGAA 3189
Qy 3192 GAGGTAGGAGGAGCTCCGAGTGTCTGGCTTCGAGAGAAAGAGAGGAGCTTATGCCGAAA 3251
Db 3188 AGACCAAGAGAGGAGGTGAAGCGTCCGGTCTGAAGAAAAGAGAGCGGATGTAGTGTGAAA 3129

Qy 3252 GGACCCGTGCAGGAAGCTGGAGCTGAGCATCTTGCACAGGGCTCTGAGACTGGACAGGCT 3311
Db 3128 GTAGATGCTCAGGAGGCAAAAACCTGAGCCCTTTTACACAAGGGAAGTGGTGGGCAGACC 3069
Qy 3312 ACTCCAGAGAGCCCTTG---AAGTTCTCTGAAGTCAAGGCAAGATGTAGAC-----CAT 3359
Db 3068 ACCCCAGAAAAGCTTTGAAAAGGCTCTCTCAAGTCAAGAGAGCATAGAGTCCAGTGAAGCTT 3009
Qy 3360 GTCCCCACGTCGCA-----GGTTATCAAGCTCCAGCAGCTCATGNA 3401
Db 3008 GTAACCACTGTCAAGCGGAAACCTTAGCTGGGGTAAATCAACAGGAGATGGTATGAA 2949
Qy 3402 CAGCCCTGGCCCTGAGTCAATCCGAAACCTTGCACACAGTGTGACAAAATGGGAAGCACT 3461
Db 2948 CAGGCTATCCCCCTGACTCGGTGGAACCCCTACACACAGTGTGATGTGAAGCACC 2889
Qy 3462 CCCTTACAGATTCAGACACTGACAGATGGGACACAGCAAGATGAACCAATGTACAGCAG 3521
Db 2888 CCGTAGCCGACTTGAACGCAACCCAGGCAACCCAGAAAAGACGAGATTGTGGAATCCAT 2829
Qy 3522 GACAGTAAAGCACTGTCAGCTGTCCAGCAGTCAAGGTCACAGAAAGAGGCGGCTACT 3581
Db 2828 GAGGAGATGAGTCCGATCTGGTACCAGTCAAGGGGACAGAAAGCAGAGGCGAGTTCCT 2769
Qy 3582 GCTCAGAAAGAGGAGCCCTTCGACACTACCTAATAATGTTCCAGCCCAAGGAACATGGG 3641
Db 2768 GCACAGAAAGAGAGCCCTCCAGCACCTTCCAGTTTGTGTTCCAGGAAGAACTAAAGAA 2709
Qy 3642 GAAGAACGGAAGAGATGTTCTTGAACCTTACACAGCAAGAGCTTACTGTGACGCGTG 3701
Db 2708 CAATCAAGATGGAAGACACTTAGAGCATACAGATAAGAGGTGTCAAGTGAACCTGTA 2649
Qy 3702 CCGTTCCTGCAAGACTGAGGTGGGTCAAGAGGTGA-----GGTTGACTGGTTG 3752
Db 2648 TCCATTCTGTCAAGACTGAGGGGACTCAAGAGGCTGACCAGTATGCTGATGAAGAAACC 2589
Qy 3753 GATGGAGAAAAGTCAAAAGAAAGACAGAGGTGTTTGTACACTCTGG-----ACCCAAC 3806
Db 2588 AAGACGTACCAATTTTCGAAGACTGTGAGGGTCTATAGACACAGGCATACAGTCACT 2529
Qy 3807 AGTCAAAAGCTGTGATGTGACATATGACAGTGAAGTGTGGAGTGGCGGGTGTGAG 3866
Db 2528 CGGGAAAAGTCACTGAAGTTGCCCTTAAAGGTGAAGGACAGAAAGCTGAATGTAAA 2469
Qy 3867 GAAAAGGAGAGTACTGAAGTGCAGAG-----TCTTAGCTGGAGGAGGA 3911
Db 2468 AAGGATGATGCTCTTGAACCTGCAGAGTCAAGCTAAGTCTCTCCATCCCCCGTGGAGAGA 2409
Qy 3912 GAGATGAAAACCTGACGTTGAAAAGGAGAAAAGGAGACAAAGCCAGAGCAAGTGAAGTGA 3971
Db 2408 GAGATGGTAGTTCAAGTTCGAAAGGGAGAAAACAGAGCAGAGCCACCCATGTGAATGAA 2349
Qy 3972 GAAGG---TGAGCAGGAAACAGCCGCTCTGTAGCATGAAGAAACCTACGGGAAGCCAGTC 4028
Db 2348 GAGAAAGCTTGAGCACGAAAACAGCTGTATCCGTATCTGAAAGGTTCAGTAAGCAGCTCTC 2289
Qy 4029 CTGACACTTGACATGCCAGCTCAGAGGGGGAAGGCCTGGAAGCCTTGGAGGAGC 4088
Db 2288 CAGACAGTGAATGTGCCCATCATAGATGGGGCAAGGAAGTCAAGCAGTTTGGGAAGAAC 2229
Qy 4089 CTTCTC---TCCAGACCAAGACAAAGCAGGTTGTCATAGAGGTTCAAGTTCAAGGCTG 4145
Db 2228 CTTCTCTCCCTCCTAGTCAAGAGGAGGCAAGTATGCCAAAATTTCAAGTTCAAGCTCT 2169
Qy 4146 GACACAAAGTCACTCAACAGCAGAAAGCTGTGGAATA-----GGTCATAGAAAACGGTT 4199
Db 2168 GAGGCATCATTTCACTCTAACAGCGGCTGCAGAGGAGAAAAGTCTTAGGAGAACTGCC 2109
Qy 4200 GTGATTTTCAAGAGCAGGTGAAGTCCAGAGTGTGTAGGTGCAACATTTATACAGCTGAG 4259
Db 2108 AACATTTTAGAAAACAGGTGAACGTTTGGAGCCTGCAAGTGTGACATTTAGTTCTTGGGAAG 2049


```

US-60-324-185-30580
; Sequence 30580, Application US/60324185
; GENERAL INFORMATION:
; APPLICANT: Morris, MacDonald
; APPLICANT: Lal, Preeti
; APPLICANT: Disp, Dinh
; TITLE OF INVENTION: METHOD FOR THE IDENTIFICATION OF SEQUENCE POLYMORPHISMS USING
; TITLE OF INVENTION: POLYNUCLEOTIDE SEQUENCE DATABASES, AND SINGLE NUCLEOTIDE
; TITLE OF INVENTION: POLYMORPHISMS IDENTIFIED THEREBY
; FILE REFERENCE: GX-0019-1 P
; CURRENT APPLICATION NUMBER: US/60/324,185
; CURRENT FILING DATE: 2001-09-21
; NUMBER OF SEQ ID NOS: 35862
; SOFTWARE: PERL Program
; SEQ ID NO 30580
; LENGTH: 6632
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 445786.20
US-60-324-185-30580

```

Query Match 35.3%; Score 2172.8; DB 76; Length 6632;
Best Local Similarity 64.4%; Pred. No. 0;
Matches 4220; Conservative 0; Mismatches 1797; Indels 539; Gaps 43;

Qy	3	GGCCGGCGGGAGTGAAGAGCCACTAGCCCATGGCGCAGGCAAGTTCCACCGAGCAGCGG	62
Db	190	GGCTTAGCGCGGGAGAAGTGCGGAGGAGCCATGGCGCGGGAGCTCCACCGAGCAGCGC	249
Qy	63	AGCCCCAGCAGCGCG---GGGGGAGCGACACGCCGAGCGAGCTGGTGTCTAGTGGCAAT	119
Db	250	AGCCCGGAGCAGCGCCCGGAGGGAGCTCCACGCCGCTGAGCCCGAGCCACGCGCGCGC	309
Qy	120	GGGCCC-----GCAGCTGAAGCCTCGGAGCAGCAGCTGGAGACCCCGCCGACCGGA	169
Db	310	GGCCCCCTCGGCCGAGGGCGCCAGACACACCGCGGACCCCGCCCATCGCTGCCTCGGA	369
Qy	170	CCCGGCCACAAGCTCCCA CAGAAAGATGCGCAGCTGTCTTCTGTCAACGGCGTAGCTGA	229
Db	370	CCCCGCCACCAAGCTCTCAGAGAAGATGGTCAAGTGTCCACCATCAATGGCGTAGCTGA	429
Qy	230	ACAAAGAGATGTCCTATCTCCAGAGGAAACACGAGGG-----	268
Db	430	GCAAGATGAGCTCAGCCTTCAGGAGGGTGACCTAAATGGCCAGAAAGGAGCCCTGAA	489
Qy	269	-----GCAGGAGGAAGAAGTCTGTTGATGAGGATGTTGGACAGCG	307
Db	490	TCAAGGAGCCCTAAACAGCCAGCGAGGAGAGGAAGTCAATGTGCACAGAGTTGGA	549
Qy	308	AGAGTCAGAGATGTGAGAGAAAAGACCGAGTTGAAGAAATGGCGGCCAACTCCACAGC	367
Db	550	AGACTCTGAAGATGTGAGCAAAAGAGACTCCGATAAAGAGATGGCTACTAAGTCAGCG	609
Qy	368	TGTTGAAGATATCACAAGAGATGGCAGGAGGAGACATCAGAAATAATTGAACAGATCC	427
Db	610	TGTTCAAGCATCACAGATGATGGGCAGGAGGAGACACCCGAATAATCGAACAGATTCC	669
Qy	428	TGCTTCAGAAAACAATGTGGAAGAAATGGTACAGCCTGCTGAGTCCAGGCTAAATGATGT	487
Db	670	TTCTTCAGAAAGCAATTTAGAGAGCTAACACAACCCACTGAGTCCCGGCTAAATGATAT	729
Qy	488	TGGCTTCAAGAAAGTATTAAATTTGTTGGTTTTAAATTCACGGTGAAGAGGATAAAA	547
Db	730	TGGATTTTAAAGAGGTGTTAAGTTTGTGGCTTTAAATTCCTGTCAAAAAGGATAAGAC	789
Qy	548	TGAAAGTTCAGATACTGTCCAACTACTACTGTCAAGAGGATGAAGGCCAAGGGSCAGA	607
Db	790	AGAGAAGCCTGCACACTGTCAGCTACTCACTGTGAAGAAAGATGAAGGGAGGGAGCAGC	849
Qy	608	AGCCTCTGTGGAGCTGGAGACCAACGAGGAGCCAGTGTGGAGACTGCGCTCGGAGAGTC	667

Db	850	A-----GGGCGTGGCGACACACAGAGCCCCAGCCTTGGGGCT-----GGAGAAGC	894
Qy	668	AGCATCCAAAGAAAGTGCAGTGAAGCAATTCACAGAGAAGCAAGAAGGCACCTCTGAAGCA	727
Db	895	AGCATCCAAAGAAAGCGAACCCAAACATCTACAGAGAAACCCGAGAGAGACCTTGAAGCG	954
Qy	728	AGAAACAGAGCAGCACAGAAATCCCTTTCAAGCCGAATCTGTATCAAGCGGCTGAGGAAGA	787
Db	955	TGACAAAGCCACGCGAAMATTTCTCCCGACGCCGAATCTGGCCAA---GCAGTGGAGGA	1011
Qy	788	AGCCAAAGATGAAGGAGAGAGAGAAACACAGAGAAAGAGCCCAAGTCCCAGAAATCCCC	847
Db	1012	ATGCAAGAGAGAAAGGAGAGAGAAACAGAGAAAGAAACCTAGCAAGTCTGCAGAAATCTCC	1071
Qy	848	GAGCAGCCCACTCAAACAGTGAGACAACTCTTCTTCAAGAAAGTTCTTCACTCAACGGTTG	907
Db	1072	GACTAGTCCCGTGACCAAGTGAACAGGATCAACCTTCAAAAAATCTTCACTCAAGTTG	1131
Qy	908	GGCCGGTGGCGAAGAACACCAAGCTTCAAGAAATCAAAGAGGATGATCTGSAACATGCG	967
Db	1132	GGCCGGCTGGCGAAAAAGAACCACTTTTCAGGAAGCCGAAGGAGATGAAGTGGAAAGCTTC	1191
Qy	968	CGAAGAGAGAAAGGACGACGAGCGACAGAAAAAGTAGACGAGAGAGAAAGAGAAACAGAGA	1027
Db	1192	AGAGAGAGAAAGAGAAACAGAGCCAGAAAAAGTAGACACAGAAAGAACGGAAGAGGACAGA	1251
Qy	1028	GCCAGCCTCGGAGGA-----GCAGGAGCCGCGACAGAGA	1060
Db	1252	GGTTGCCTCCGAGAAACTGACCGCTCCGAGCAAGCCACCCACAGAGAGCCGCGAGAAG	1311
Qy	1061	CACAGACGAGCCAGGTTGTCAGCAGACTACGAGAAGGTGGAGCTGCCTTTTGGAAAGACCA	1120
Db	1312	TGCCACGAGCCCCGGTTATCAGCTGAATATGAGAAAGTTGAGCTGCCCTCAGAGGAGCA	1371
Qy	1121	GGTTGGTGACCTGGAGGCATCGTCAGAGAGAAAGTGTGCTCTCTTTGGCAACGGAAAGTGT	1180
Db	1372	AGTCAGTGGGCTCGCAGGGACCTTCTGAAGAGAAAACTGCTCGTTGGCGACAGAAAGTGT	1431
Qy	1181	TGATGACAAGATGGAAAGCCACCCAGAA---GTTCTTCAGAGCTTCCACGTGACACCGT	1237
Db	1432	TGATGAGAAATAGAAGTCCACAGAGAGAGGTTGTGGCCGAAGTGCACGTGACGACCCGT	1491
Qy	1238	GGAGAAGACAGAGAGGAGCAGGGAGGAGGAGAGGCTGAAGGGGCGTGGTGGTAGA	1297
Db	1492	GGAGAGAGAACCCAGAGACAGAAA-----ACGAGAGTGGGA	1527
Qy	1298	AGGAACAGGAGAAATCTCTTGCCCTTGAGAAATGCTGCTGAGCCCCAGGAGTCCCCCAGGA	1357
Db	1528	AGAAACAGCAGGGTCTGTGCCAGCTGAAGAAATGGTTGAAATGGATGACAGAACTTCAGGA	1587
Qy	1358	AGCTGAGCCTGCTGAGAGCTGATGAAGAGCAGAGAGATGTGTCTCTCTGAGGAGGACCA	1417
Db	1588	AGCCGAACCTTGCCAGAGGCTGGTGAAGCTCAAGAGAAACGTTGTTTCCGGAGAGGAGACC	1647
Qy	1418	CACCTCAACTGACAGACTAAGTCTCTGAAGAGAAAGCGTGCCTCCAAAACACCCAGAGGCGAT	1477
Db	1648	TACACAGGAGCTGACCTCAGTCTGTAGAAAGGTGTCTGTCTCTGTCTGTCTGTCTGTCTGT	1707
Qy	1478	TGTCAGTGAGGTGGAGATGCTGTCTCTCAGGAAAGAAATCAAGGTACAGGGAAGTCTCCCT	1537
Db	1708	TGTGAGTGTAGGTGGAAATGCTGTCTATCAAGAGAGAGATGAAGGTGCGAGGGAAGTCCACT	1767
Qy	1538	GAACAACTCTTCAGTGTCTCAGGCTTAAAGAGCTGTCTGGAGAGAGCAGAGAGGGAA	1597
Db	1768	AAAGAAGCTTTTACAGCAGCTGGCTTAAAAAAGCTTTCTGGAAGAGAAACAGAAAGGGAA	1827
Qy	1598	ACGAGGAGGTGGGGAGACGAAGAGCCTGGAGAAATACCAACATTCACACCGAATCCCC	1657
Db	1828	AAGAGGA---GGAGGAGAGAGAGAAATCAGGGAGGACCACTCAGGTTTCAGCCGATTTCTCC	1884
Qy	1658	AGAGAGTGTGATGACAGAGGGAGAGAGCTCTCGCTCGTCCCCCGAGGAGCCTGAGGA	1717
Db	1885	GGACAGCCAGGAGGAGCAAAAGGGCGAGAGCTCTCTCTATCTCTCTCTCTCTCTCTCTCTCT	1944

QY 1718 GACCACTGCTGAGAAAGGCGCTGGAAGACCCAGGATGGGGAAAGCTGGAAGG 1777
 Db 1945 GATCACTGTCTGGAAAAGGGCTTAGCCGAGGTGACAGAGATGGGGAAGCTGAAAGG 2004
 QY 1778 AACATCTCCGATGGAGAAAGAAAGAGATCACTCCCTGGGCACTCTTCAAAA 1837
 Db 2005 AGCTACTTCCATGGAGAAAGAAAGAAAGAGTGTCACTCCCTGGGCACTCTTCAAAA 2064
 QY 1838 GATGTGACACCCAAAGAAAGCGGTCGGAAGACCTTCTGAGATGACAAAGAGAAAGCT 1897
 Db 2065 GATGTGACGCCCAAGAGCGCTTAGACGCGCTTCCGAAAGTGAATAAGAGATGAGCT 2124
 QY 1898 GGAGAAAGCTCAAGAGCCCACTTGTCTCCCTGATAGCAAGTGTCAAGAAATGCAAGA 1957
 Db 2125 GGACAAAGGTCAAGAGCGCTACTTGTCTTCCACGAGAGCAAGCGCTTGAATATCAGA 2184
 QY 1958 TGAAGTCAAAACCTGTGGTGAAGAACAAAAGCCAGAGAACCAAGCGTAGGGTAGTAC 2017
 Db 2185 AGAAATGAAAGGAGCGTGGAAAGCCAAAGCCGAAAGAACCAAGCGTAGGTATC 2244
 QY 2018 TTCACTGCTTGGGAAAGCACTGATTTGTGTGATCATCAAGAAAGAGCAAGAGGC 2077
 Db 2245 CTCACTATCTTGGGAAGCTTTATTTGTGGATCATCAAGAAAGAGCAAGAGGAG 2304
 QY 2078 ATCTCTTCAATGATGAAGAGGCGCAAGCACTGGAGGGGACGTCAAGAGCA 2137
 Db 2305 GTCTCTTCTGATGAAGAGGGGGACCAAAAGCAATGGAGAGACCAACAGAAAGCTGA 2364
 QY 2138 GGAAGCCCAAGAAAGCAAGAACCCGGAACAGACGCTGTTCCGCGCAGACCCAGAGCA 2197
 Db 2365 TGAGGCGCGGAAAGACAAAGAGACGGGAGACAGCGGATCTTGTGTTCCAGAAAG 2424
 QY 2188 GGACCAAGGCAAGAGAGTCTCTCAACCCGAGCAGCGGAAAGCCCTTCCGAAAGGAG 2257
 Db 2425 TGAATCAAGGCAAGAGAGTCTCTCCCGAGCAGCTGGAAGCCCTTACGAAAGGAGG 2484
 QY 2258 TGTCTCACTTGGAGTCAATTTAAAGATTAGTCACTCAAGAAATAATCCAGTCAA 2317
 Db 2485 CATTCTCACTGGAGTCAATTTAAAGATTAGTCAAGCAAGAAATAATCCAGTCAA 2544
 QY 2318 ACTGGAAGAAAGCCGGAAGAC-----TCTAGTATGAGACAGTGTCCACTGA 2365
 Db 2545 GCTGGAAGAAAGGAAAGCTCCATAGCTGGGTCTGGTGTGAACATTTCCACTCAAG 2604
 QY 2366 GATCGAAGCCGATGAGAGAAATCTTGGTTTCCATTAAAGAAATTCATCCCGAGCGG 2425
 Db 2605 CACTGAAGCCGTAAGAAAGAAATCTGGGTCTCAATCAAGAAATTTATCTGAGCAAG 2664
 QY 2426 GAAAGAAAGGCGCAGACGGGAAAGCAAGCAAGCTGTGGAAAGTCAAGGGCACTGGA 2485
 Db 2665 GAAAGAAAGGCGCAGATGGGAAAGCAAGCAAGCCCTGTTGAAAGCGCAGGGGCAAG 2724
 QY 2486 GATTAATGAGAGACCCATAGTGTCCAGCGGTGCTGTCTGTGATTAATGATGAGT 2545
 Db 2725 GGGCAAGAGATGATGATGTATCTCCGCGGTGCTGTCTGTGATTAATGATGAGT 2784
 QY 2546 GGAGAGGAGAAAGATGGAAGCCAG-----GGAAATAGAGAGTCCCGAGCTGC 2595
 Db 2785 AGAAAGGAGAAATGGAAGCAAGCAAGCCCAAAAGCCGAGAGAGCCCGAGAGAG 2844
 QY 2596 TGGGGCTGTGTATGATGTCTGAGAGAGTCAATAGACTGTGTCTGATGATGATG 2655
 Db 2845 AGGCAAGCACTAGGTGTCTCAAGAGGTCAAGAGTCAAGTTCATATGATGAGCAG 2904
 QY 2656 CAGTCAATGATGGGACAGGGCAGTCAAGCTGTGCAAGAGCGGTCTCTCTGATGAT 2715
 Db 2905 CTGTCTGTGAGGAGAGAGGAGCTCAATTTATGAAAGAGGTCTCTCTTGTGATAT 2964
 QY 2716 CGGCTTCCGTAAGAGACCTTGAACACACAGCGGAGAGGCAATGCCACTGTTGAAG 2775
 Db 2965 CTGCTTCACTGACAGAACTTCTTGAACAGTAAAGCTGAAGCCGCACTGTTAACGAG 3024

QY 2776 AGCTACTGAAAGAAACATATTC---AGAAAGAACTCTGTGTCTCAAGCCAGATTC 2832
 Db 3025 AGGTATTGAAAGAAAGTATTTGACAGAAAGAACCCCAAGCTTAACTGAACCTCTGC 3084
 QY 2833 CAGAGGGTAAAGATGCCATGACAGCATGTGTCACTAGTGAATGATTTCACTCAGAG 2892
 Db 3085 CAGAGAAACAGAGAGCGCGGGGCGACAGCGTCTTATGTAGGCGGAAATGACCCCGAAG 3144
 QY 2893 CTGTGACAGCCACAGAGACCTCAGAGGCTCCGTACTGAAAGATTACGGAAGCATCG 2952
 Db 3145 CTGTGACAGCTGCAGAAACTGACAGGCGCATTTGGGTGCGAAAGAAAGCCAGAGATCTG 3204
 QY 2953 GGCGCGAAGAGACCAAGACATGTGTGTCCGAGTCTTCCAGCTGACTGTCCCAAGCA 3012
 Db 3205 CTGTGAAAGAACCAAGAAATGATGTAGCAGCTCCCAATTAACGAGCTCCCAAGCA 3264
 QY 3013 CCAAGAGAAAGCCACCCCACTTCAAGAGTGAAGATGTGTGTCTGATTAAGAGAG 3072
 Db 3265 CCACAGAGAGGCGCATCTCGGTGAGAGGTGAAAGGTGGGTACTTGAATAGAGAGC 3324
 QY 3073 AGAGCGCGACACCGAGCGCATCTCCCAAGCGCTTCCAGCAAGGTGAAAGAGAGTCC 3132
 Db 3325 AAGAGAGCGGACCTCAAGAGTCTTCCAGGAGTGGCAAGAAAGTGAAGAGAGATCC 3384
 QY 3133 AGGTGCTGTCAAGC---CAGACTGTGACAGAGCGGGTCAAAAGCACTGAGAGAGTTG 3189
 Db 3385 AGCTGCTGTGACCGGTGGGCGAAGATGTGCTTCAAGCTGTGACAGACAGAGCGAG 3444
 QY 3190 AAGAGTGAAGAGAGACTCCGAAGTGTGCTTGGAGAAAGAGAGAGCTTATGCCGA 3249
 Db 3445 AAAGACAGAAAGAGCAGGCTGAAGGCTGGGTCTGAAGAAAGAGAGCGATGTATGTTGA 3504
 QY 3250 AAGAGCCCGTGCAGGAAGCTGAGTGAACATTTGCAAGGCGCTGTGAGACTGAGCAG 3309
 Db 3505 AAGTATGATCTCAGAGAGCAAAACCTGATTTTACCAAGGAGAGGTGTGGGCGAGA 3564
 QY 3310 CTACTCCAGAGAGCCTTG---AAGTCTGAAAGTCAAGCGAGATGTAGAC-----C 3357
 Db 3565 CCAAGCCAGAAAGCTTTGAAAGAGTCTTCAAGTCAAGAGAGCTTAAAGTCAAGTGC 3624
 QY 3358 ATGTGCGCAAGTGC---GTTATCAAGCTCAAGAGCTGATGG 3399
 Db 3625 TTGTAACTGTCAAGCCGAACTTAGCTGGGTAAATTCAGAGAGATGTGATGG 3684
 QY 3400 AACAGCCCGTGGCCCTGAGTATCCGAAACCTTGAACAGACAGTGAAGCAATGGAAGA 3459
 Db 3685 AACAGGCTATCCCTCTGATGTGGAAGCCCTTACAGACAGTGAAGTGAAGCA 3744
 QY 3460 CTCCCTTGAAGATTCAGACATGAGATGGGACACAGCAAGATGAACATTTGAACAGC 3519
 Db 3745 CCCCCGTAGCCGACTTTGACGACAGAGCAACACCAAGAAAGAGATTTGTGAAATCC 3804
 QY 3520 AGGACAGTAAAGCCACTGAGCTGTAGGCACTCAAGTCAAGAAAGAGCGGCTA 3579
 Db 3805 ATGAGAGAAATGAGGTGATCTGTGACCAAGTCAAGGGGCGCAGAAAGAGGAGCTTC 3864
 QY 3580 CTGTCTGAAGAAAGAGCTTGAACATCACTAATTAATTTTCAAGCCAGAGAAACATG 3639
 Db 3865 CTGCAAGAAAGAGAGGCTTCAGACCTTCAAGTGTGTTCAGAGAAAGCTAAAG 3924
 QY 3640 GGGAGAAACAGAGAGAGATGTTCTTGAACCTTACAGCAAGAGCTTACTGTGAGCGG 3699
 Db 3925 AACATCAAGATGAGAGACCTTGAAGATACAGATTAAGAGGTGTCAAGTGAAGCTG 3984
 QY 3700 TGCCCGTTTGGCAAGACTGAGGTGGGTCAAGAGGTGA-----GTTGACTGCT 3750
 Db 3985 TATTCATTTGTCAAGAGTGAAGGAGCTCAAGAGGTGACAGTATGTGATGAGAAAG 4044
 QY 3751 TGAATGAGAAAGAGTCAAGAAAGAAAGAGAGGTGTTGTAACCTGTG-----ACCA 3804
 Db 4045 CCAAGAGCGTACATTTTTCAGAGAGCTTGAAGGAGCTTATTAACACAGGCAATACAGTGA 4104
 QY 3805 ACAGTCAAAAGCGTGTGATGTGACATATGACAGTGAAGTGAAGTGGAGCGGCTGTC 3864

Db 4105 GTCCGGAAAGGTCACCTGAAGTTCCTTAAAGGTGAAGGACAGAAGAGCTGAATGTA 4164
QY 3865 AGGAAAGAGAGATGATGAAGTGCAGAG-----TCTTAGCTGGAGAGG 3909
Db 4165 AAAAGGATGATGCTCTTGAACCTGCAGAGTCACGCTAAGTCTCTCCCTCCATCCCGCTGGAGA 4224
QY 3910 GAGAGATGGAACCTGACGTTGAAAGAGGAAAGGAGACAAAGCCAGAGCAAGTGAAGT 3969
Db 4225 GAGAGATGGTATGTTCAAGTTCGAAAGGAGAAACAGAAAGCAGAGCCAAACCATGTGAATG 4284
QY 3970 AAGAAGG---TGAGCAGGAAACAGCCGCTCTGAGCATGAAGGAACCTACGGGAAGCCAG 4026
Db 4285 AAGAGAGCTTGAGCAGCAAGACAGCTGTACCGTATCTGAAGAGTCAAGAGAGCTCC 4344
QY 4027 TCTGACATCTGACATGCCAGCTCAGAGAGGGGAAAGGCACTGGGAAGCTTGGAGGAA 4086
Db 4345 TCCAGACAGTGAATGTGCCATCATAGATGGGCAAGGAAGTCAGCAGTTTGGAGGAA 4404
QY 4087 GCCCTTCTC---TCCAGACCAAGACAAAGCAGGTTTGCATAGAGTTCAAGTTCAAAGCC 4143
Db 4405 GCCCTCTCCCTGCTAGGTCAGAGAGGAGGAGTATGCACCAAAATTCAGATTCAGAGCT 4464
QY 4144 TGGACACAAAGTCACCTCAAAACAGCAGAGCTGTGGAAGA-----GCTCATGAAGACGG 4197
Db 4465 CTGAGGATCATCTCTTAACAGCGGCTGCAGAGGAGAAAGGCTCTTAGGAGAAACTG 4524
QY 4198 TTGTGATTTTCAGAGACAGGTGAAAGTCCAGAGTGTGTAGGTGCACTTATTACCAGCTG 4257
Db 4525 CCAACATTTTGAAGAACAGGTGAAACGTTGGAGCTGCAGGTCGACATTTAGTTCTGGAAG 4584
QY 4258 AAGAGTCTCTGCAACCGGTGGCCACTGGACTCTTACAGATGCAGAGACACGGTACCCC 4317
Db 4585 AGAAATCTCTGAAGAAATGAAGACTTGCCTCATCCAGGGGAAAGATGCTGTGCCCA 4644
QY 4318 TGGGGCTGAGTCTCAGGACAGATCCATCCCAATCATAGTAACCTCTGCTCTGAAAGCA 4377
Db 4645 CAGGCCCGACTGTCAGGCAAAATCGACACAGGTGATAGTATCTGCTACTACCAAGAAAG 4704
QY 4378 CCTTACATCTCGACCTCAAGGAGAAATAAGCGCATCCAGAGAGAGCGATCAGAGGAAG 4437
Db 4705 GCTTAAGTTCGACCTCGAAGGAGAGAAACCAATCATCACTCAAGTGAAGTCAGATGAAG 4764
QY 4438 AGGACAGCCAGATGCTGTGCTCTGATGCTGACGCAAGGAGAGTACAGCAATCGAAAAG 4497
Db 4765 TCGATGACGAGGTTGCTTGCC-----AGGAGGTCAAAGTCAAGTGTAGCAATTCAGGAGG 4818
QY 4498 TCCTCAAGGCTGAACCTGAGATCTGGAACCTGAGAGTGAAGCAACAAGATTGTGCTGA 4557
Db 4819 ATTTAGAGCTGAAATGGGAATTTTGGAACTTGAGACCAAAAGCAGTAAACTTGTCCAAA 4878
QY 4558 ACCTCATTCAGACAGCCGTTGACCAAGTTGCAAGTAC---AGAAACAGCCCGGAAACTC 4614
Db 4879 ACATCATCCAGACAGCCGTTGACCAAGTTTGTAGCTAAGAAAGAACAGCCAGCAATGT 4938
QY 4615 ATGCTTATGATTCAGACAGCCAGGTTTCT-----CTGCATGAGGCTTTGAC 4658
Db 4939 TGAGCTGTGATTTACAGACACAAGCTCACGTGATAAAAGCTGACAGCCAGGAGCTGGAC 4998
QY 4659 AGCAGGAGCCCAACAGATGCTGCACAAAATGAAGATGCCAAGATGAACACCCAGTG 4718
Db 4999 AGGAAACGGAAGAAAGGAGAGGAACTCAGAGGCTCTGACAGGATGAACACCAATTA 5058
QY 4719 CCGCAGCCCA-----GAGAGGACTTGCAGAGTCTGACCGTTCTG 4757
Db 5059 CTTTCAGCCAAAGAGGAGTCAGAGTCAACCGCAGCTGGGACCAAGCATCTCTGATATTTCCA 5118
QY 4758 GAGGCATGGGCTCAGCCTCGGAAATGCTTGCAGGCTTGCAGTTGAAAGCCCGGCTGCA 4817
Db 5119 AAGACATGAGTGAAGCCTCAGAAAAGACCATGACTGTTGAGGTAGAAGGTTCCACTGTA 5178
QY 4818 A----- 4818

Db 5179 ATGATCAGCAGCTGGAAGAGGTGCTCTCCATCTGAGGAGAGGGAGGTGGAGCTGGAA 5238
QY 4819 -----AGTAAGCATTTGAGAAGCTGC 4838
Db 5239 CAAAAGTCTGTGCAGAAAGATGATGCTCATGCTTGTGTAGCAGAAAGAAATAGAGAAGTCA 5298
QY 4839 CTCTCAACCCCAAGA--TCCAAAAGGAGCATGCTGCTGATGCGCCCTCAGCTCCAAGCT 4896
Db 5299 TAGTTGAACCCGAAAGAGATGAAGAAAGGTGATGATGTTGATGACCCCTGAAAAACGAACT 5358
QY 4897 TAGCCAGGAGCAGAGGCGCAGTCCCTTGGAAACCTTAACCAAGAAATCCCCAGACACCAACG 4956
Db 5359 CAGCCCTGGCTGATCTACTGATGCTCAGAGGCTTAACCAAGAGTCCCAGATACAAATG 5418
QY 4957 GACAAAAGCTAAACGAGGAGGCGATCCCCAAAAGTTGAGGTCAGGAAAGAAATGT 5016
Db 5419 GACAAAACAAAAGAGAGAGGAGTGCACCAAGGAGTAGAATTCAGAGAAAGGAAAGTGC 5478
QY 5017 CTACCAAGTCAGTCAAGAGAAACAGGCCCGCAGGACAGAGAGGACCTGCAGAGGCAAGG 5076
Db 5479 ACAGTGAATCAGATAAAGCGATCACACCCCAAGCAGCAGGAGAGTTACAGAAAACAGAGA 5538
QY 5077 GAGACCTGGCAGAACTCTAAGATGTTAGTTG-----CTCATTTGT 5115
Db 5539 GAGATCTGCAAGTCAGAACTTACAGAACTTTAAACATCATGCGAGTTAACTCATTTGT 5598
QY 5116 ACATCTGTAAGACAGAAATGTGAAAAACAAGTCACAGAACAA-----GATGCTGCTGTTGGG 5171
Db 5599 CTGTTTGGAAAGACAGAAATGTGAAGCAAGTAGTAGAAGAAATGAATGCTGCTGCTGAG 5658
QY 5172 ACCTTGAGACCAAGATTTTCAGAGCCCATGAGATCCAGAGAGAGGCGCTCCAAATGATTT 5231
Db 5659 A-CTGAAGACCAAGTATTTTCAGA--ACTTTGAGAAATGAGAGCAGGACATCAACTGATCT 5716
QY 5232 CCACCAGTAGAGCACCCCGACAAATCTGAGGCTTCATCGGAGCTAGAGCCAGCTCAACA 5291
Db 5717 CATTTCTAGAGAGC-CCCTGACAACTCTGAGGCTTCATCAGGAGCTAGAGCCATTTAAACA 5775
QY 5292 TTTCTCGTTTCAAGACTGCTTTGATTTGCCCTTTGATGCCGCTGATTTCTTAACA 5351
Db 5776 TTTCTCTTTTCCAAGACCAACCTACAATTTTCCCTTGATAAC----- 5817
QY 5352 TTTCTCGTTTCAAGACTGCTTTGATTTGCCCTTTGATGCCGCTGATTTTCGGATTT 5411
Db 5818 -----CATATAAATTTCTGATTT 5833
QY 5412 TAAGTCTCGGTTCTCAACCTCGAAC-CAATTTCTGCCATACCTAGTTCCACTTCTCAAA 5470
Db 5834 TAAGGTCTTAAATTTCTTAACCTCGAACTGGAGTTGGCAATACCTAGTTCTGCTTTGAAA 5893
QY 5471 CTGGAGCATCCTCTTTATGATTTATATGATGATGTTTATGATGCTCTCTCTGTAACCT 5530
Db 5894 CTGGAGTATCATTTCTTTACATATTTATATGATGTTTAAAGTAG---TCTCTCTGATCT 5950
QY 5531 ATTGTATATTTTTTCTTAACGTTTT-----AAGCACATGCTTTTGTAT 5573
Db 5951 ATTGTATATTTTTTCTTAATGTTTAAAGGAAATGTCAGGATACTACATGCTTTTGTAT 6010
QY 5574 TATGCAATATATAACGGGTGTCAGCCATAGCAGCTTTGAAAAGCTCCAAGCCTCAAC 5633
Db 6011 CACACAGTATATATGATGGGCAATGTCATAGTCAGGCTTTGAGGAGCTTTAAAGCCTCAGT 6070
QY 5634 TGTAACTCGAGCAAAACAGATAACATTC-----CTGGCAAGAGAGACAAGTCTTT 5684
Db 6071 TATATAACCCACAAAAACAGAGCTCTCTAGATGTAACATTCCTGATCAAGGTACAATTC 6130
QY 5685 TTTAAAGTTTATGATGCTTTAGATCTGTGGGCTTCTAGTCTCTCTGAAAGTGGTGTTC 5744
Db 6131 TTTAAAAATTCACCTAATGATTTAGGTCATATTTAGTGGTACTCTGAAATTTGGTCACTTTC 6190
QY 5745 CTATGACAGCAGCTCAGAAATAAAAAACCCATTTTGAACATCCAGGATGTCACCAATA 5804
Db 6191 CTATTACAGGAGTGTGCTAATAAATCTAAAAAGCATTTTGAACACATACAGAAATGTTCTATTG 6250

```

QY 5805 TTACCATGATTTTTCCTCCCTTTTGGCAATCCAGTCCAGTGGAAAGTCTCCTC 5864
Db 6251 TCATTGGGAAATTT-----TTCTTCTAACCCAGTGGAGGTAGAAAGATATATT 6303
QY 5865 TGTGTCAGA-----TTAAGCCCTGCTCTTAATGATATGACAAATGA 5907
Db 6304 CTGGTAGCAATTAACCTTACATCTTTTCTACTTGTATGTTGTTGACCGATTA 6363
QY 5908 GTGTGCTTAAGGCCATGAGA-----TGTTCCTAATGCAGAAAGAAATCT 5951
Db 6364 GTGTGCTTAATCTGAGGCAAGTAGTAATATGTTTATATGTTATGAAAGAAAT 6423
QY 5952 GTTGTACGTTTTTTGATGTACTCTCTGATGCTGACCGAATGCATATGCAGATGAA 6011
Db 6424 GTTGTAAAG--TTTTGATTTTACTCTTAATATGCTGACTGCATTCACATGCAATGAAA 6481
QY 6012 TGAGTCTGTTCTTTACAGATGATATTTGATAGATACAGAGTTGCTGTGTATATC 6071
Db 6482 TAAGTCAAGTTCTTTACCAATGATATTTGATAGATACGATGTGTGTGTCCATATT 6541
QY 6072 TGTGCCCCCTTCTTTAAGAACAAATGTCATTAATGTTCTTTGATTAATTTGATTTGAC 6131
Db 6542 TGTGCCATTTCTTTAAGAACAAATGTTGCAACACATTCATTGATTAAGTTGTGATTTGAC 6601
QY 6132 AACTGATTTAAATAAA 6147
Db 6602 GACTGATTTAAATAAA 6617

```

Search completed: December 13, 2002, 00:00:20
 Job time : 11603 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 12, 2002, 14:19:40 ; Search time 391 Seconds
(without alignments)
11377.814 Million cell updates/sec

Title: US-09-902-432-3
Perfect score: 6160
Sequence: 1 cgggcccgggagtagaag.....aaataaacatttgactac 6160

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 416730 seqs, 361097816 residues

Total number of hits satisfying chosen parameters: 833460

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents,NA New:*

- 1: /cgn2_6/ptodata/2/pna/PCT_NEW_COMB.seq:*
- 2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq:*
- 3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq:*
- 4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq:*
- 5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:*
- 6: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq:*
- 7: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5204.2	84.5	5236	6	US-10-152-319A-2063
2	2144.4	34.8	6608	1	PCT-US02-04915-43
3	1994.8	32.4	5931	5	US-09-724-676-12735
4	1994.8	32.4	5931	5	US-09-724-676A-12735
5	1920.8	31.2	5773	5	US-09-724-676-12736
6	1920.8	31.2	5773	5	US-09-724-676A-12736
C 7	550	8.9	563	6	US-10-152-319A-1312
8	162.2	2.6	178	6	US-10-152-319A-2137
9	124.6	2.0	180	6	US-10-152-319A-1193
10	119	1.9	753	5	US-09-724-676-12734
11	119	1.9	753	5	US-09-724-676A-12734
12	71	1.2	575	6	US-10-203-138A-9167
13	71	1.2	1969	6	US-10-203-138A-4036
14	70.8	1.1	1926	6	US-10-194-046-3
C 15	70.8	1.1	8705	6	US-10-273-678-16
C 16	70.8	1.1	8705	6	US-10-291-230-14
C 17	70.8	1.1	8705	6	US-10-291-249-14
C 18	70.8	1.1	16080	5	US-09-724-566A-48
19	68.4	1.1	107	5	US-09-620-607B-815
20	66.4	1.1	3489	6	US-10-194-046-1
C 21	65.8	1.1	659158	5	US-09-771-208A-20
22	61	1.0	700	6	US-10-203-138A-5915
23	56.4	0.9	2550	5	US-09-724-676-41735
24	56.4	0.9	2550	5	US-09-724-676A-41735
25	56.4	0.9	2709	5	US-09-724-676-41736
26	56.4	0.9	2709	5	US-09-724-676A-41736

27 56.4 0.9 2805 5 US-09-724-676-41728 Sequence 41728, A
28 56.4 0.9 2805 5 US-09-724-676A-41728 Sequence 41728, A
29 56.4 0.9 2964 5 US-09-724-676-41729 Sequence 41729, A
30 56.4 0.9 2964 5 US-09-724-676A-41729 Sequence 41729, A
31 56.4 0.9 3382 5 US-09-724-676-41723 Sequence 41723, A
32 56.4 0.9 3382 5 US-09-724-676A-41723 Sequence 41723, A
33 56.4 0.9 3637 5 US-09-724-676-41717 Sequence 41717, A
34 56.4 0.9 3637 5 US-09-724-676A-41717 Sequence 41717, A
35 56.4 0.9 3817 5 US-09-724-676-41745 Sequence 41745, A
36 56.4 0.9 3817 5 US-09-724-676A-41745 Sequence 41745, A
37 56.4 0.9 4072 5 US-09-724-676-41732 Sequence 41732, A
38 56.4 0.9 4072 5 US-09-724-676A-41732 Sequence 41732, A
39 56.2 0.9 474 6 US-10-203-138A-3499 Sequence 3499, Ap
40 56.2 0.9 906 6 US-10-273-334-33 Sequence 33, Appl
41 55.4 0.9 2142 6 US-10-152-319A-1333 Sequence 1533, Ap
42 55 0.9 905 6 US-10-273-334-11 Sequence 11, Appl
43 55 0.9 907 6 US-10-273-334-9 Sequence 9, Appl
44 55 0.9 907 6 US-10-273-334-13 Sequence 13, Appl
45 55 0.9 907 6 US-10-273-334-25 Sequence 25, Appl

ALIGNMENTS

RESULT 1

US-10-152-319A-2063
; Sequence 2063, Application US/10152319A
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Higgins, Brandon
; APPLICANT: Castile, Arthur
; APPLICANT: Elashoff, Michael
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5089-US
; CURRENT APPLICATION NUMBER: US/10/152,319A
; CURRENT FILING DATE: 2002-05-22
; PRIOR APPLICATION NUMBER: US 60/292,335
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/297,523
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,925
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,810
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/303,807
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/303,808
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/315,047
; PRIOR FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: US 60/324,928
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/330,867
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: US 60/330,462
; PRIOR FILING DATE: 2001-10-22
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2221
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2063
; LENGTH: 5236
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. NM_057103
US-10-152-319A-2063

Query Match 84.5%; Score 5204.2; DB 6; Length 5236;
Best Local Similarity 99.7%; Pred.No. 0;
Matches 5233; Conservative 0; Mismatches 13; Indels 1; Gaps 1;

QY 1 CGGGCCGGGAGTGAAGAGCCCTGAGCCATGGGGCGAGGCTTCCACCGAGCAGC 60
 Db 1 CGGGCCGGGAGTGAAGAGCCCTGAGCCATGGGGCGAGGCTTCCACCGAGCAGC 60
 QY 61 GGAGCCCCGAGCAGCCGGCGGAGACGACACGCCGAGCAGCTGTGTCTCAGTGGCCATG 120
 Db 61 GGAGCCCCGAGCAGCCGGCGGAGACGACACGCCGAGCAGCTGTGTCTCAGTGGCCATG 120
 QY 121 GGCCCGCAGCTGAAGCCTCGGAGCAGCTGAGACCCCGCGACGCGGACCCCGCACCA 180
 Db 121 GGCCCGCAGCTGAAGCCTCGGAGCAGCTGAGACCCCGCGACGCGGACCCCGCACCA 180
 QY 181 AGCTCCCAAGAAATGGCCAGCTGTCTTCTGTCAACGGCTTACTGAAACAAGAGATG 240
 Db 181 AGCTCCCAAGAAATGGCCAGCTGTCTTCTGTCAACGGCTTACTGAAACAAGAGATG 240
 QY 241 TCCATGTCCAAAGAGAAACCGAGAGGGCGAGAGAAAGTCTTGTATGAGATGTTG 300
 Db 241 TCCATGTCCAAAGAGAAACCGAGAGGGCGAGAGAAAGTCTTGTATGAGATGTTG 300
 QY 301 GACACGAGAGTCAGAAAGATGTGAGAGAAAGAACCGAGTTGAGAAATGGCGGCAACT 360
 Db 301 GACACGAGAGTCAGAAAGATGTGAGAGAAAGAACCGAGTTGAGAAATGGCGGCAACT 360
 QY 361 CCACAGCTGTGAAATATCAAAAAGATGGCGAGAGAGACATCAGAAATATTTGAAC 420
 Db 361 CCACAGCTGTGAAATATCAAAAAGATGGCGAGAGAGACATCAGAAATATTTGAAC 420
 QY 421 AGATCCCTGTCCAGAAACAAATGTGAGAAAGATGTACAGCTGTGATCCAGGCTA 480
 Db 421 AGATCCCTGTCCAGAAACAAATGTGAGAAAGATGTACAGCTGTGATCCAGGCTA 480
 QY 481 ATGATGTGGCTTCAAGAAAGTATTTAAATTTGTTGTTTAAATTCACGATGAAGAG 540
 Db 481 ATGATGTGGCTTCAAGAAAGTATTTAAATTTGTTGTTTAAATTCACGATGAAGAG 540
 QY 541 ATAAATATGAATATGATGATCTGTCTCACTGTCAAGAAAGATGAAGGCGAAG 600
 Db 541 ATAAATATGAATATGATGATCTGTCTCACTGTCAAGAAAGATGAAGGCGAAG 600
 QY 601 GGGCGAAGGCTCTGTCCGAGCTGAGACCAACAGAGGCCAGTGTGAGACTGCGCTG 660
 Db 601 GGGCGAAGGCTCTGTCCGAGCTGAGACCAACAGAGGCCAGTGTGAGACTGCGCTG 660
 QY 661 GAGATCAGCATCCAAAGAAAGTGAAGCAATCCACAGAGAACCAAGAGGACCC 720
 Db 661 GAGATCAGCATCCAAAGAAAGTGAAGCAATCCACAGAGAACCAAGAGGACCC 720
 QY 721 TGAAGCAAGACAGAGCAGCAAGAAATCCCTTCAAGCCGAAATCTGATCAAGCGGCTG 780
 Db 721 TGAAGCAAGACAGAGCAGCAAGAAATCCCTTCAAGCCGAAATCTGATCAAGCGGCTG 780
 QY 781 AGGAAGAAAGCCAAATGAAGAGAAAGAAACAGAGAAAGGCCACCAAGTCCCGAC 840
 Db 781 AGGAAGAAAGCCAAATGAAGAGAAAGAAACAGAGAAAGGCCACCAAGTCCCGAC 840
 QY 841 AATCCCCGAGCAGCCAGTCAACAGTGAACAACATCTTCTTCAAGAGTTCTTCACTC 900
 Db 841 AATCCCCGAGCAGCCAGTCAACAGTGAACAACATCTTCTTCAAGAGTTCTTCACTC 900
 QY 901 ACGGTTGGGCGCGCTGGCCAGAGAACCAAGCTTCAAGAAATCAAAAGGAGTGAATCG 960
 Db 901 ACGGTTGGGCGCGCTGGCCAGAGAACCAAGCTTCAAGAAATCAAAAGGAGTGAATCG 960
 QY 961 AAACTGGCCGAGAGAAAGAGAGCAAGGCAAGAAATAGCGAGAGAAAGAGAA 1020
 Db 961 AAACTGGCCGAGAGAAAGAGAGCAAGGCAAGAAATAGCGAGAGAAAGAGAA 1020
 QY 1021 AGACAGAGCAGCCTTGAAGAGCAGAGAGCCGCGCAGAGACACAGACAGGCGCTTGT 1080
 Db 1021 AGACAGAGCAGCCTTGAAGAGCAGAGAGCCGCGCAGAGACACAGACAGGCGCTTGT 1080
 QY 1081 CAGCAGACTTACAGAGAAAGGTGAGCTGCTTTGGAGAACCAAGGTTGATCTGAGGAGCAT 1140

Db 1081 CAGCAGACTTACAGAGAAAGGTGAGCTGCTTTGGAGAACCAAGTTGTGTGATCGAGGCAT 1140
 QY 1141 GGTCAAGAGAGAAAGTGTCTCTTTGGCAAACGAAAGTTTGTATGAGAAATGAAACCC 1200
 Db 1141 GGTCAAGAGAGAAAGTGTGTCTTTGGCAAACGAAAGTTTGTATGAGAAATGAAACCC 1200
 QY 1201 ACCAAGAAAGTTTGTGAGAGGTCCACTGAGCAACCTGTGAGAAAGACAGAGAGAGCAG 1260
 Db 1201 ACCAAGAAAGTTTGTGAGAGGTCCACTGAGCAACCTGTGAGAAAGACAGAGAGAGCAG 1260
 QY 1261 GAGAGAGAGAGAGAGGCTGAAGGGGCGTGTGTGAGAGAAACAGAGAAATCTTGCCTCC 1320
 Db 1261 GAGAGAGAGAGAGAGGCTGAAGGGGCGTGTGTGAGAGAAACAGAGAAATCTTGCCTCC 1320
 QY 1321 CTGAGAAACCTGGCTGAGCCCGAGAGGTCCTCCAGAAAGCTGAGCTGAGAGGCTGA 1380
 Db 1321 CTGAGAAACCTGGCTGAGCCCGAGAGGTCCTCCAGAAAGCTGAGCTGAGAGGCTGA 1380
 QY 1381 TGAAGAGCAGAGAAATGTGTCTCTGAGAGACCACTCACTCACTGACAGACCTTAAGTC 1440
 Db 1381 TGAAGAGCAGAGAAATGTGTCTCTGAGAGACCACTCACTCACTGACAGACCTTAAGTC 1440
 QY 1441 CTGAAGAGAAAGCCTGCGCCCAACACCCAGAAAGCATTTGTCAGTGAGAGATGCTGT 1500
 Db 1441 CTGAAGAGAAAGCCTGCGCCCAACACCCAGAAAGCATTTGTCAGTGAGAGATGCTGT 1500
 QY 1501 CCTCTCAGAGAAAGATCAAGGTATCAGGAAAGTCCCTTGAAGAACTCTTCAATGATGCTAG 1560
 Db 1501 CCTCTCAGAGAAAGATCAAGGTATCAGGAAAGTCCCTTGAAGAACTCTTCAATGATGCTAG 1560
 QY 1561 GCTTAAAGAGCTGTCTGGAGAAAGCAAGAGGGGAAACAGAGAGGTGGGAGACAGAG 1620
 Db 1561 GCTTAAAGAGCTGTCTGGAGAAAGCAAGAGGGGAAACAGAGAGGTGGGAGACAGAG 1620
 QY 1621 AGCCGTGAGAAATCAACACATTCACACCCGAAATCCCGAGAGAGTGTGATGAGAGAG 1680
 Db 1621 AGCCGTGAGAAATCAACACATTCACACCCGAAATCCCGAGAGAGTGTGATGAGAGAG 1680
 QY 1681 GAGAGAGCTGTGCTGCTGCTCCCGAGAGAGCTGAGAGACCAAGTGTCTGAGAGAAAGGC 1740
 Db 1681 GAGAGAGCTGTGCTGCTGCTCCCGAGAGAGCTGAGAGACCAAGTGTCTGAGAGAAAGGC 1740
 QY 1741 CGCTGAGAGCAGCCCAAGATGGGAGAGCTGAGAGAGAACTTCCGATGAGAGAGAA 1800
 Db 1741 CGCTGAGAGCAGCCCAAGATGGGAGAGCTGAGAGAGAACTTCCGATGAGAGAGAA 1800
 QY 1801 AGAGAGAGGAGTCACTCCCGGAGATCCTTCAAAAGATGGTGAACCCCAAGAAACGGG 1860
 Db 1801 AGAGAGAGGAGTCACTCCCGGAGATCCTTCAAAAGATGGTGAACCCCAAGAAACGGG 1860
 QY 1861 TCCGAGAGCTTCTGAGAGTGAAGAGAGAGAGCTGAGAGAGTCAAGAGCCGCACT 1920
 Db 1861 TCCGAGAGCTTCTGAGAGTGAAGAGAGAGAGCTGAGAGAGTCAAGAGCCGCACT 1920
 QY 1921 TGTCTCACTGATGACACAGTGTGAGAAATGCAAGATGAAGTCAAACTGTGTGAGG 1980
 Db 1921 TGTCTCACTGATGACACAGTGTGAGAAATGCAAGATGAAGTCAAACTGTGTGAGG 1980
 QY 1981 AACAAAGCCGAGAGAAACCAAGGCTAGGGTGAATCTTCAAGTCTTGGAGACACTGA 2040
 Db 1981 AACAAAGCCGAGAGAAACCAAGGCTAGGGTGAATCTTCAAGTCTTGGAGACACTGA 2040
 QY 2041 TTTGTGTGATTCATCCAAAGAGAGCAAGGCAATCTCTTCAAGATGATGAAGAG 2100
 Db 2041 TTTGTGTGATTCATCCAAAGAGAGCAAGGCAATCTCTTCAAGATGATGAAGAG 2100
 QY 2101 GGCCAGAGACACTGGAGGGGACGTCAAGAGCAGAGAGGCCAGCAAGACAAAGAG 2160
 Db 2101 GGCCAGAGACACTGGAGGGGACGTCAAGAGCAGAGAGGCCAGCAAGACAAAGAG 2160
 QY 2161 CCGGAAAGAGAGCTTCTTCTGAGCAACCCAGAGAGCAAGGCGCAAGAAAGTTCT 2220

Oy	4381	TATTCCTGACCTTCAAGAGAAATAATTAAGCGATCCGAGAGAGGCCATCGAGAGAAAGAG	4440
Db	4381	TATATCTCGACCTTCAAGAGAAATAAGCCGATCCGAGAGAGGCCATCGAGAGAAAGAG	4440
Oy	4441	ACAAGCCAGATGCTGATCTTGATGCTCTGACGCAAGAGAGATGACGAAATGAAAAAGTCC	4500
Db	4441	ACAAGCCAGATGCTGATCTTGATGCTCTGACGCAAGAGAGATGACGAAATGAAAAAGTCC	4500
Oy	4501	TCGAAGCTGAACCTTGAGATCTTGGAACCTTGAGAGTTAAGAGCAACAGATTGTGCTGAACG	4560
Db	4501	TCGAAGCTGAACCTTGAGATCTTGGAACCTTGAGAGTTAAGAGCAACAGATTGTGCTGAACG	4560
Oy	4561	TCATTCAACAGACCGTTGACCAAGTTCCGACGTAAGAAAACAGCCCCGGAATCATGCTT	4620
Db	4561	TCATTCAACAGACCGTTGACCAAGTTCCGACGTAAGAAAACAGCCCCGGAATCATGCTT	4620
Oy	4621	ATGATTCACAGACCAGGTTCCCTCATGCAAGGCTTGAACAGCAGAGAACCCMAACAGATGCT	4680
Db	4621	ATGATTCACAGACCAGGTTCCCTCATGCAAGGCTTGAACAGCAGAGAACCCMAACAGATGCT	4680
Oy	4681	GGACAAAAATGAAAAGATGCCAAGATGAAACACCCAGTCCGACGCCCAAGAGAGACTTGC	4740
Db	4681	GGACAAAAATGAAAAGATGCCAAGATGAAACACCCAGTCCGACGCCCAAGAGAGACTTGC	4740
Oy	4741	AAGTCCGACCGTTCTGGAGGCAATGGGCTCAGCCTCGGAAATGCTTGCGCGCTTGACAGT	4800
Db	4741	AAGTCCGACCGTTCTGGAGGCAATGGGCTCAGCCTCGGAAATGCTTGCGCGCTTGACAGT	4800
Oy	4801	TGAAAGCCCGGCTGCAAAAGTAACATTTGAGAAAGCTGCTCTCAACCCAAAGTCCAAA	4860
Db	4801	TGAAAGCCCGGCTGCAAAAGTAACATTTGAGAAAGCTGCTCTCAACCCAAAGTCCAAA	4860
Oy	4861	AGGAGCATGCTGCTGATGCGCTTCAAGCTCCAAAGCTTAAGCCCAAGCAGAGGCCAGTGCT	4920
Db	4861	AGGAGCATGCTGCTGATGCGCTTCAAGCTCCAAAGCTTAAGCCCAAGCAGAGGCCAGTGCT	4920
Oy	4921	CTGGAACCTTAACCAAAATAATCCCAACACCAAGGCAACCAAGGCTTAACCGAGAGAGGGC	4980
Db	4921	CTGGAACCTTAACCAAAATAATCCCAACACCAAGGCAACCAAGGCTTAACCGAGAGAGGGC	4980
Oy	4981	ATCCCCCAAAAGTTGAGGCTCCAGAGAGAAATGTTACCAAGTCAGTCAAAAGAGAACAA	5040
Db	4981	ATCCCCCAAAAGTTGAGGCTCCAGAGAGAAATGTTACCAAGTCAGTCAAAAGAGAACAA	5040
Oy	5041	AGGCCAGGCAAGAGGACCTGCAAGAGCCAAAGGAGACCTTGCGCAAAATCCTTAAGATG	5100
Db	5041	AGGCCAGGCAAGAGGACCTGCAAGAGCCAAAGGAGACCTTGCGCAAAATCCTTAAGATG	5100
Oy	5101	TTAGTTGCTCATTTGACATCTGTGAAGCCAGAAATGTAAAAAAMAATCAAGAACCAAGATG	5160
Db	5101	TTAGTTGCTCATTTGACATCTGTGAAGCCAGAAATGTAAAAAAMAATCAAGAACCAAGATG	5160
Oy	5161	CTGCTGTTGGGACCTTGAGACCAAGATTTTCAGAGCCCATGAGATCCAGAGAGAGGGCCG	5220
Db	5161	CTGCTGTTGGGACCTTGAGACCAAGATTTTCAGAGCCCATGAGATCCAGAGAGAGGGCCG	5220
Oy	5221	TCCAATGATTTCCACC	5237
Db	5221	TCCAATGATTTCCACC	5237
RESULT 2			
PCT-US02-04915-43			
: Sequence 43, Application PC/TUS0204915			
: GENERAL INFORMATION:			
: APPLICANT: Murray, Richard			
: APPLICANT: Glynn, Richard			
: APPLICANT: Watson, Susan R.			
: APPLICANT: Aziz, Natasha			
: APPLICANT: Eos Biotechnology, Inc.			
: TITLE OF INVENTION: Methods of Diagnosis of Angiogenesis, Compositions and			
: FILE REFERENCE: 018501-006200PC			

Query Match	Best Local Similarity	Score	DB 1	Length
Matches 4215; Conservative	64.3%	Pred. No. 0;	Mismatches 1796; Indels 543; Gaps 45	
3	GGCCGGCGGAGTAGAAGACCACTGAGCCATGGCGCAGCGACGATTCCACCGACGACGG	62		
162	GGCTAGGGCGGGAGAAAGTGGCGAGAGCCATAGGGCGCCGGAGAGCTCCACCGACGACCGC	221		
63	AGCCCCGAGCAGCCG---GGGGGAGGCAACGCGGAGGAGCTGCTGCTCAGTGGCCAT	119		
222	AGCCCGGAGCGCCCGGAGGGAGCTCCACGCGGCTGAGCCGAGCCGACGCGCGCG	281		
120	GGGCGCCGAGCTGAAGCCTCGGGA-----GCAGCTGAGAGCCCGCCGACGCGGAC	170		
282	GGCCCTCGGGCGGAGGCGGCGGCGACACACGCGGAGCCCGGCATGCTCCTCGGAC	341		
171	CCCGCCACCAAGCTCCACAGAAAGATGGCCAGCTGTCTTCTGTCAACGGCTGATGA	230		
342	CCCGCCACCAAGCTCTCAGAAAGATGGTCAAGCTGTCCACCATCAATGGCGTAGCTGAG	401		
231	CAAGAGATGTCATGTCCAGAGAGAAAACCGAGAGG-----	268		
402	CAAGATGAGCTAGCCTCCAGAGGGGTGACCTAAATGGCCAGAAAGAGCCTGAACGGT	461		
269	-----GCAGAGGAGAAAGTCTGTGATGAGGATGTTGACACGCA	308		
462	CAAGAGGCGCTTAACAGCGACGAGGAGAAAGATCATTTGTCAACGAGGTTGACAGANA	521		
309	GAGTCAGAAAGATGTGAGAGAAAAGACCGAGTTGAAGAAATGGCGGCACTTCACAGCT	368		
522	GACTCTGAAGATGTAGGAGGAAAGACATCCGATTAAGAGATGGCTAATAGTCAGCGGTT	581		
369	GTTGAAGATATCAAAAGGATGGGACGAGGAGCATCAGAAATATTAACAGATCCCT	428		
562	GTTCAAGCATACAGATGATGGGACGAGGAGGA---ACCGAATATTCGAACAGATTCT	638		
429	GCTTCAGAAAACATGTGGAGAAATGTATACGCTGTGATGCCAGGCTAATGATTT	488		
639	TCTTCAGAAAGCAATTTGAAAGAGCTAACACACCACTGATCCAGGCTAATGATTT	698		
489	GGCTTCAAGAAAGTATTTAAATTTGTGGTTTAAATTCACGATGAGAGAAATTAAT	548		
699	GGATTTAAGAAAGTGTAAATTTGTGGCTTTAAATTCATGTGAAAAAGGATAAGACA	758		
549	GAAAGTCAGATACCTGTCCAATCTACTCTCAAGAGAGATGAAAGGCGAAGGGCGAGAA	608		
759	GAGAAAGCTGACACTGTCCAGCTACTCTGTGAAGAAAGATGAAGGGAGGAGGACAGCA	818		
609	GCCTCTGTGCGAGCTGGAGACCAACGAGGAGCCCAAGTGTGAGATACGCTCGGAGAGTCA	668		
819	-----GGGGCTGGCGACCAACAGAGCCCAAGCTTTGGGGCT-----GGAGAAACA	863		

QY 669 GCATCCAAAGAAAGTGAAGTGAAGCAATCCACAGAGAAAGCAAGGACACCTTGAAGCAA 728
Db 864 GCATCCAAAGAAAGCGAAACCCAAACAATCTACAGAGAAACCCGAAGAGACCTTGAAGCGT 923
QY 729 GAACAGAGCAGCAGAGAAATCCCTTTCAGAGCCGAATCTGATCAAGCGGCTGAGGAGAA 788
Db 924 GAGCAAGCCAGCAGAGAAATTTCTCCCGAGCCGAAATCTTGCCAGAGCAGT---GGAGAA 980
QY 789 GCCAAGATGAAGAGAGAGAGAAACAAAGAGAAAGAGCCACCAAGTCCCGCAGATCCCG 848
Db 981 TGCAAGAGAGAGAGAGAGAGAAACAAAGAAAGAAACCTAGCAAGTCTGCAGATCTCCG 1040
QY 849 AGAGCCAGTCAACAGTGAAGCAACATCTTCTTCAAGAGTCTTCACTCAGCGTGG 908
Db 1041 ACTAGTCCCGTGACAGTGAACAGAGTCAACCTTCAAAATAATCTTCACTCAAGGTGG 1100
QY 909 GCCGGCTGGCGCAAGAGACCAAGCTTCAAGAAATCAAAAGAGGATGATCTGGAACCTGCC 968
Db 1101 GCCGGCTGGCGCAAAAGACCAAGTTTCAGNAGCCGAGAGGATGAAGTGAAGCTTCA 1160
QY 969 GAGAAGAAAGGAGCAGAGGCGCAGAAAAGTAGACAGAGGAAGAAAAGGAAAGACAGAG 1028
Db 1161 GAGAAGAAAGGAAACAAGAGCCAGAGAAAAGTAGACACAGAGAAAGACGAAAGGCAGAG 1220
QY 1029 CCAGCCTCGAGGA-----GCGAGAGCCGCGCAGAGAC 1061
Db 1221 GTTGCCTCCGAGAAATGTAGCCCTCCGAGCAAGCCACCCACAGGAGCCGCGCAGAAAGT 1280
QY 1062 ACAGACCAGGCGAGTTGTGACAGACATACAGAGAAAGTGAGGCTGCTTTTGAAGACCAG 1121
Db 1281 GCCACAGAGCCCGTTATCAGCTGAATATAGAGAAAGTTGAGCTGCCCTCAGAGAGCAA 1340
QY 1122 GTTGGTGACCTGGAGGAGTGTGACAGAGAGAAAGTGTCTCTCTTGGCAACGGAAGTGT 1181
Db 1341 GTCAGTGGCTCGCAGGACCTTCTGAAGAGAAACCTGCTCGTGGTGCGACAGAGTGT 1400
QY 1182 GATGAGAAGATGGAAGCCACCAAGAA---GTGTTGAGAGGTCACGTGAGCACCGTG 1238
Db 1401 GATGAGAAATAGAAAGTCCACCAAGAGAGAGTTGTGCGCAAGTCCACGTGAGCACCGTG 1460
QY 1239 GAGAAGACAGAGGAGGAGGAGGAGGAGGAGGAGGCTGAGGGGGCGTGGGTAGAA 1298
Db 1461 GAGGAGAAACCGAAGACAGAA-----ACGGAGTGGAA 1496
QY 1299 GGAACAGAGAAATCTTGCCTCCCTGAGAACTGGCTGAGCCCGCAGAGGTCCTCCAGGAA 1358
Db 1497 GAACACAGGGTCTGTGCCAGCTGAAGATTTGTTGGAATGATGACAGAACCTCAGGAA 1556
QY 1359 GCTGAGCCTGTGAGGAGCTGATGAAGAGCAGAGAGATGTGTCTCTGAGGAGAGCCAC 1418
Db 1557 GCCGAACCTGCAAGGAGCTGGTGAAGCTCAAGAAACGCTGTGTTCCGGGAGAGGCCCT 1616
QY 1419 ACTCAACTGACAGACCTTAAGTCTGAGAGAGAGAGCTGCCCAACACCCAGAGGCAAT 1478
Db 1617 ACACAGGAGCTGACCTCAGTCTGATGAGAGAGTGTGTTCAAAACCCCGCAGAGGCGTT 1676
QY 1479 GTCAGTGAAGTGAAGTGTCTCTCTCAGGAAAGATCAAGGTACAGGGAAGTCCCTTG 1538
Db 1677 GTGAGTGAAGTGAAGTGTCTATCACAGGAGAGATGAAGTGAAGGAGTCTCACTA 1736
QY 1539 AAGAACTCTTCAAGTAGCTCAGGCTTAAAGAGCTGTCTGGGAAGAGAGCAAGGGGAAA 1598
Db 1737 AAGAAAGCTTTTACCAAGCACTGGCTTAAAGAGCTTCTGGAAGAAACAGAAAGGAAA 1796
QY 1599 CGAGGAGGTGGGAGAGAGAGAGCTGGAGAAATACCAACATTCACCCGAATCCCCA 1658
Db 1797 AGAGGA---GGAGAGAGCAGAGGAATCAGGGAGACACACTCAGGTTCCAGCCGATCTCCG 1853
QY 1659 GAGAGTGTGATGACAGAGAGGAGAGAGCTCTCGCTGCTCCCGAGGAGCCCTGAGGAG 1718
Db 1854 GACAGCAGAGAGGAGGAGAGGAGAGCTCTGCTCTATCCCTGAGGAGCCCGAGGAG 1913

QY 1719 ACCAGTGTCTGAGAAAGGGCCGCTGGAAAGCAACCCAGAGTGGGGAAGCTGAGGAAGGA 1778
Db 1914 ATCACGTGTCTGAAAGGGCTTTAGCCGAGGTGCAGCAGATGGGGAAGCTGGAAGGA 1973
QY 1779 ACTACTTCCGATCGAGAGAAAGAGAGAGGATCACTCCCTGGGCATCTCTTCAAAAAG 1838
Db 1974 GCTACTTCCGATCGAGAGAAAGAGAGAGGTGTCACTCCCTGGGCATCTTCAAAAAG 2033
QY 1839 ATGGTGACACCCCAAGAAACGGGTCCGAAGACTTCTGAGAGTCAACAGGAGGAGAGCTG 1898
Db 2034 ATGGTGACGCCCAAGAAAGCGTGTAGACGCCCTTCGAAAGTGTAAAGAAAGATGAGCTG 2093
QY 1899 GAGAAAGTCAAGAGCGCCACCTTGTCTCCACTGATAGCAGTGTGAGAAATGCAAGAT 1958
Db 2094 GACAAAGTCAAGAGCGCTACCTTGTCTTCCACCGAGACAGCCTCTGAAATGCAAGAA 2153
QY 1959 GAAGTCAAAAATCTGTGTGAGGAAACAAAGCCAGAGAGCAAAAGCGTAGGGTGATACT 2018
Db 2154 GAATGAAAGGAGCGTGGNAGAGCCAAAGCCGAGAGCAAAAGCCAGAGTGGATACC 2213
QY 2019 TCAGTGTCTTTGGAAGCACTGATTTGTGTGCGATCATCTCAAGAGAGAGCAAGGAAGCA 2078
Db 2214 TCAGTATCTTTGGAAGCTTTAAATTTGTGTGGATCATCTCAAGAAAGAGAGCAAGGAAG 2273
QY 2079 TCCTCTTCAGATGATGAAGAGGGCCNAGGACACTGGAGGGGACAGTTCACAGAGCAGAG 2138
Db 2274 TCCTCTTCTGATGAGGAAGGGGACCAAAAGCAATGGAGGAGACCAACCAAGAAAGCTGAT 2333
QY 2139 GAGCCAGCAACAGCAAAAGCGGAAACAGACGCTTCTCCTCCAGACCCAGGAGCAG 2198
Db 2334 GAGCCGGAAGAAAGCAAGAGACGGGACAGCGGATCTTCTGTTCCCAAGAAACAT 2393
QY 2199 GACCAAGCGCAAGAAAGTCTCTCACCCGAGCCAGCGGAGACCCCTTCCGAAGGGGAAGGT 2258
Db 2394 GATCCAGGCGAGGAAAGTTCTCTCCCGAGCAAGCTGGAAGCCCTACCGAAGGGAGGC 2453
QY 2259 GTCTCACTTTGGAGTCATTTAAAGATTAGTCACTCCAAAGAAAAAATCAAGTCAAAA 2318
Db 2454 GTTTCACCTCTGGAGTCATTTAAAGGTTAGTCACGCCAAGAAAAAATCAAAAGTCCAAG 2513
QY 2319 CTGGAAGAGAAACCGCAAGAC-----TCTAGTGTAGAGAGTGTCCACTGAG 2366
Db 2514 CTGGAAGAGAAAGCGAAGACTCCATAGTGGGTCTGGTGTAGAACATTCATCTCCAGAC 2573
QY 2367 ATCGAACCCGATAGAGAAAGATCTTGGGTTTCCATTAAGAAATTCATCCCGGACGCGG 2426
Db 2574 ACTGAACCCGTTAAAGAGATCTCTGGTCTCAATCAAGAGTTTATCTCGACGAGG 2633
QY 2427 AAGAAAGGCGACAGCGGAAAGCAAGCAACAGCCACTGTGGAAGACTCAGGGCCAGTGGAG 2486
Db 2634 AAGAAAGGCGCAGATGGGAAACAAAGAACAGCCCTGTTGAAGACGCGAGGCCAACAGGG 2693
QY 2487 ATAAATGAGGACACCTTAATGTCCAGCGCTGCTGCTGTCTGAGTATATGCAAGTG 2546
Db 2694 GCCAAGCAAGATGACTGTGATGTCCCGGCGTGGTCCCTCTGTCTGAGTATGATGTGTA 2753
QY 2547 GAGAGGAGAGAAAGTGG-----AAGCCAGGGGAATACGGAGGCTGCCCGAGCTGCTG 2597
Db 2754 GAAAGGAGAAATGAGGCGACAGCAAGCCCAAAAGGCGCAGCAGCCCGAGCAGAG 2813
QY 2598 GGGGCTGTGTACTGTCTCGAGGAGCTCAGTAAGACTCTGGTCCACTGTGAGTGTGCA 2657
Db 2814 GCAGCCACTGAGGTGTCCAAGGAGCTCAGCGAGAGTCAAGTTTATATGATGAGCAGCAGCT 2873
QY 2658 GTCAATTGATGGGACCGGCGAGTCAACAGTGTGAGAGCGGTCTCTCTGTTGATATCC 2717
Db 2874 GTCGCTGACGGGACGAGGCGAGCTACCAATTATTGAAGAAAGGTCTCTCTCTGATATCT 2933
QY 2718 GCTTCCGTAAACAGAACTCTTGAACACAGCGGAGAGCCATGCCACCTGTTTGAAGAG 2777
Db 2934 GCTTCAGTGAAGAACTCTTGAACAGTGAAGCTGAAGCCGCACTGTTAACTGAGGAG 2993
QY 2778 GTCACCTGAAAAAGACATCATTCG---AGAAGAAACTCTGTGTCTCACCAGACGTTACCA 2834

QY 4819 -----AGTAAGCATGTGAGAGCTGCCT 4840
Db 5205 AAGTCTGTGCAGAGATGATGTCATGCTTGTAGCAGAAAGAAATAGAGAAGTCACTA 5264
QY 4841 CCTCAACCCCAAGA--TCCAAAAGGAGCATGCTGCTGATGCGCCCTCAGCTCCAAAGCTTA 4898
Db 5265 GTTGAACCGAAGAGATGAAGAAGGTGATGTTGATGACCCCTGAAACCGAACTCA 5324
QY 4899 GCCAGCAGAGGCCAGTGCCTCTGGAACCTTAACCAAGAATCCCCAGACACCAACGGA 4958
Db 5325 GCCCTGGCTGATCTGATGCTCAGGAGGCTTAACCAAGAGTCCCAGATACAAATGGA 5384
QY 4959 CCAAAGTAAACGAGGAGGGATCCCCAAAAGTTGAGGTCCAGGAAGAAATGTCT 5018
Db 5385 CCAAAACAAAAGAGAGAGGAGTGCACAGGAAGTAGAATTGCAGGAAGGAAAGTGCAC 5444
QY 5019 ACCAAGTCAGTCAAGAGAGAACAGGCCAGCAGAGAGGACCTGCAGGAGCCAAAGGGA 5078
Db 5445 AGTGAATCAGATTAAGCGATCACACCCAAAGCAGAGGAGTTACGAAGAACAGAGAGA 5504
QY 5079 GACTGGCAGAACTCTAAGATGTTAGTTC-----CTCATTTGTAC 5117
Db 5505 GAATCTGCAAGTCAGAACTTACAGATCTTAAACATCATGAGTTAAACTCATTTCT 5564
QY 5118 ATCTGAAGACCAAGATGTGAAGAACAAAGTCACAGAAC--GATGCTGCTGTTGGAC 5173
Db 5565 GTTTGGAAGACCAAGATGTGAAGACAAAGTAGTAGAAGAAATGAATGCTGCTGAGAG- 5623
QY 5174 CTTGAGACCAAGATTTTCAGAGCCATCAGATCCAGAGAGCAGGCGCTCCAAATTTCC 5233
Db 5624 CTGAAGACCAAGTATTTTCAGAG-ATTTTGAAGATTTGAGAGCAGGACATCAACTGATCTCA 5682
QY 5234 ACCAGTAGAGACCCCGCAATTTCTGAGGCTTCATCGGAGCTAGAGCCAGCTAACTT 5293
Db 5683 TTTCTAGAGAGC-CCCTGACAACTCTGAGGCTTCATCAGAGCTAGAGCCATTTAATTT 5741
QY 5294 TCCTCGTTTCAAGACTGCTTTGATTTGCTTGTGATGTCGCGTGGTATTTCTTAACATTT 5353
Db 5742 TCCTCTTTCAAGACCAACCTACAATTTTCCCTTTGATAAC----- 5781
QY 5354 TCCTCGTTTCAAGACTGCTTTGATTTGCCCCCTTGATGTCGCTGCTGATTTTCGGATTTA 5413
Db 5782 -----CATATAAATTTCTGATTTA 5799
QY 5414 AGGTCTGCTGCTTCTCAACTGGAAAC-CAATTCGCCATACCTAGTTCACACTTCTCAAACT 5472
Db 5800 AGTCTCTAATTTCTTAACCTGGAACTGGAGTTGGCAATACCTAGTTCCTGTTCTGAACT 5859
QY 5473 GGAGCATCTCTTTATGATTTATATGATGTTTATGATGATGCTCTCTCTGTAACCTAT 5532
Db 5860 GGAGTATCATTTCTTACATATTTATATGATGATGTTTAAAGTAG--TCCTCTGATCTAT 5916
QY 5533 TGATATTTTTTTCTAACGTTT-----AAGCACATGCTTTTTGTATTA 5575
Db 5917 TGATATTTTTTTCTTAAGTTTAAAGGAAATGTGCAGGATACATACATGCTTTTTGTATCA 5976
QY 5576 TGCAATATATAACGGGTGTGCAGCATAGCAGAGCTTTTGAAGAGCTCCAAAGCTCAACTG 5635
Db 5977 CACAGTATATGATGGGCATGTGCCATAGTGAGGCTTTGGGAGCTTTAAGCCTCAGTTA 6036
QY 5636 TAACTGCAGCAACAGATAAATTC-----CTGGCAAGAAGACAGCAAGTCTTTTT 5686
Db 6037 TATAACCCCAAAAAACAGAGCCTCTAGATGTAAACATTTCTGATCAAGGTACAATTTT 6096
QY 5687 TAAAGTTTACTGATGCTTAGATCTGTGGGCTTCTAGTCTCTGAAAGTGGTGTGTTTCTCT 5746
Db 6097 TAAAAATTCACATAATGATGAGGTCATATTTAGTGTGACTCTGAAATTTGCTCACTTTCTCT 6156
QY 5747 ATGCACAGCGAGCTCAGAAATAAAAAACCCCATTTTGAACATCCAGGATGCTCCCAATTT 5806
Db 6157 ATTACGGAGTGTGCCAAAACCTAAAGCAATTTTGAACATACAGAAATGTTCTATTGTC 6216

QY 5807 ACCATGATTTTTTCCCCCTTTTTCCTTAATCCAGTCCAGGTTGGAAAGAAAGTCTCTCTG 5866
Db 6217 ATTGGGAAATTTTGC-----TTTCTAACCCAGTGGAGGTTAGAAAGAAAGTTATATCT 6269
QY 5867 TGTTCAGA-----TTAAGCCCTGTCTCTTAATGATATGACAAATCAGT 5909
Db 6270 GGTAGCAATTAACCTTTTACATCCTTTTCTACTGTTTATGTTGTTGGACCGATAAGT 6329
QY 5910 GTCCCTAAGGCCCATGAGA-----TGTTTCTTAATGCAGAAAGAAATCTGT 5953
Db 6330 GTCTCTTAATCCTGAGCAAGTAGTGAATATGTTTATATGTTATGAAGAAAGAAATTTGT 6389
QY 5954 TGTACGTTTTTTGATGTTACTCTTATGCTGGACCGAAATTCATATGCAGATCGAAGTG 6013
Db 6390 TGTAAAG--TTTTTGTATCTACTCTTATATGCTGAGCTGCATTCACACATGCCATGAAATA 6447
QY 6014 AGTCCTGTTCTTTTACAGATGGTATTTTGTAGATACCTGGAGTTTGTGTTTATATCTG 6073
Db 6448 AGTCAGGTTCTTTTACAAATGGTATTTTGATAGATACCTGGATTTGTGTTGCCATATTTG 6507
QY 6074 TGCCCTCTTTTAAAGAACAAATGTTGCATTTATGTTTCTTTTGGATAAAATTTGTGATTTGACAA 6133
Db 6508 TGCCATTTCTTTTAAAGAACAAATGTTGCAACACATTCATTTGGATAAGTTGTGATTTGACGA 6567
QY 6134 CTGATTTAAATAAA 6147
Db 6568 CTGATTTAAATAAA 6581

RESULT 3
US-09-724-676-12735
; Sequence 12735, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12735
; LENGTH: 5931
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676-12735

Query Match 32.4%; Score 1994.8; DB 5; Length 5931;
Best Local Similarity 67.6%; Pred. No. 0;
Matches 3281; Conservative 0; Mismatches 1332; Indels 237; Gaps 25;

QY 3 GGCGCGCGGAGTAGAAGAGCCACTGAGCCATGGGCGCAGGAGTTCCACCGAGCAGCGG 62
Db 162 GGTAGCGCGGGAAGAGTGGCGGAGGCCATGGGCGCGGAGCTCCACCGAGCAGCGC 221
QY 63 AGCCCGAGCAGCGG---GCGGGGAGCAGACGCGCGAGGAGCTGTTGCTCAGTGGCCAT 119
Db 222 AGCCCGAGCAGCGCGCGCGAGGAGCTCCACGCGGCTGAGCCCGAGCCAGCGGCGG 281
QY 120 GGGCCCGAGCTGAAGCCTCGGGA-----GCAGTGGAGACCCCGCAGCGCGGAC 170
Db 282 GGGCCCTCGCGGAGCGCGCGCAGACACACCGCGGACCCCGCATCGTCTCGGAC 341
QY 171 CCGGCCACCAAGCTCCACAGAGAATGCCAGCTGCTCTTCTCAACCGCTAGCTGAA 230
Db 342 CCGGCCACCAAGCTCTTACAGAGAATGTGCTGCTGCTCCACCATCAATGCGGTAGTGA 401
QY 231 CAAGGAGATGTTCCATGTCCAAGAGGAAAAACAGGAGG----- 268
Db 402 CAAGATGAGCTCAGCCTCCAGGAGGGTGACTTAATGGCCAGAAAGGCCCTGAACGGT 461
QY 269 -----GCAGGAGGAAGAGTCTGTTGATGAGGATGTTGGAACGCGA 308
Db 462 CAAGGAGCCTTAAACAGCCAGGAGGAAGAAAGTCAATTGTCACAGAGGTTGGACAGAGA 521

QY 309 GAGTCAGAAATGTGAGAGAAAAGCCGAGTTGAAAGAAATGGCCGCACTTCCACAGCT 368
Db 522 GACTCTGAAATGTGAGGAAAGAGATCCGATAAAGATGGCTACTAAAGTCAACGGTT 581
QY 369 GTTGAAGATATCAAAAGATGGGAGGAGGAGCATCATGAAATTAATTGAACAGATCCCT 428
Db 582 GTTACGACATCAAGATGATGGGAGGAGGAGACATCCGAAATTAATGAAACAGATTCCT 641
QY 429 GCTTCAGAAAATGATGAGAAATGATACAGCTGCTAGTCCAGGCTAATGATGTT 488
Db 642 TCTTCAGAAATTAATGAAAGCTTAACAAACCACTAGATCCAGGCTAATGATATT 701
QY 489 GCGTTCAAGAAATTTAAATTTGTGTTTAAATTCACGATGAAGAGATTAATAAT 548
Db 702 GGATTTAAGAAAGTGTTTAATGTTGTGCTTTAAATTCAGTGAATAAAGATTAAGCA 761
QY 549 GAAAAGTCAGATCTGTCACTACTCACTGTCAAGAAAGATGAAGGCGAAGGGGAGCA 608
Db 762 GAGAAAGCTTGACATCTGTCCAGCTACTCACTGTGAAGAAAGATGAAGGGAGGAGCAGCA 821
QY 609 GCGTTGTGAGAGCTGAGAGCACCAGAGGCCAGTGTGAGACTGCGTCCGAGAGTCA 668
Db 822 -----GGGGCTGGGAGCACCAAGACCCAGCCTTGCGGCT-----GGAAGAGCA 866
QY 669 GCATTCAAAAGAAATGATGAGCTGAAGCAATTCACAGAGAAAGAAAGGCACTTGAAGCA 728
Db 867 GCATTCAAAAGAAAGGCAACCCAAACATCTACAGAGAAACCCGAAGAGACCTTGAAGCT 926
QY 729 GAAAGAGAGAGAGCAAGAAATCCCTTCAGAGCCGAAATCTGATCAAGCGGCTGAGAGAA 788
Db 927 GAGCAAGCCAGAGAAATTTCTCCCAAGCCGAAATCTGGCCAA---GCAATGAGAGAA 983
QY 789 GCCAAAGATGAAGAGAAAGAAAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 848
Db 984 TGCAAAAG 1043
QY 849 AGCAGCCAGATCAAGATGAGAGCAATCTTCTTCAAGAGATCTTCACTACGATGG 908
Db 1044 ACTATCTCCGTGACCAAGTGAACAGATCAACCTTCAAAAATTTTCACTCAAGGTTGG 1103
QY 909 GCGGCTGGCGCAAGAGCCAGCTTCAAGAAATCAAAAGAGATGATGTGAAACTGCC 968
Db 1104 GCGGCTGGCGCAAAAG 1163
QY 969 GAGAAAGAAAG 1028
Db 1164 GAGAAAGAAAG 1223
QY 1029 CCAGCTCGAGAG-----GAGAGAGCCGAGAGAGAGC 1061
Db 1224 GTTGTCTCCGAGAACTGACCGCTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1283
QY 1062 ACAGAGCCAG 1121
Db 1284 GCGCAG 1343
QY 1122 GTTGTGAGCTGAG 1181
Db 1344 GTTGTGAGCTGAG 1403
QY 1182 GATGAG 1238
Db 1404 GATGAG 1463
QY 1239 GAG 1298
Db 1464 GAG 1499
QY 1299 GGAACAG 1358
Db 1500 GAAACAG 1559

QY 1359 GCTGAGCTGTGAG 1418
Db 1560 GCCAGAGCTGCAAG 1619
QY 1419 ACTCACTGACAG 1478
Db 1620 ACAG 1679
QY 1479 GTCAGTGAAGTGAAGATGCTGCTCTCAGAGAAAGATCAAGAGAGAGAGAGAGAGAG 1538
Db 1680 GTGAGTGAAGTGAAGATGCTGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1739
QY 1539 AAGAACTCTTCACTAGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1598
Db 1740 AAGAACTCTTCACTAGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1799
QY 1599 CGAGAGAGTGGGAG 1658
Db 1800 AG 1856
QY 1659 GAGAGTGTGATGAG 1718
Db 1857 GACAGCAG 1916
QY 1719 ACCAGTGTCTGAG 1778
Db 1917 ATCAAGTGTCTGAG 1976
QY 1779 ACTACTTCCAGTGAAG 1838
Db 1977 GCTACTTCCAGTGAAG 2036
QY 1839 ATGCTGACCCAG 1898
Db 2037 ATGCTGACCCAG 2096
QY 1899 GAGAGAGTGAAG 1958
Db 2097 GAGAGAGTGAAG 2156
QY 1959 GAGTCAAACTGTTGAG 2018
Db 2157 GAATGAAG 2216
QY 2019 TCACTGTCTGAG 2078
Db 2217 TCACTGTCTGAG 2276
QY 2079 TCCCTTTCAGATGATGAAG 2138
Db 2277 TCCCTTTCAGATGATGAAG 2236
QY 2139 GAGGCGAG 2198
Db 2337 GAGGCGAG 2336
QY 2199 GACCAAGCGCAAG 2258
Db 2397 GATCCAGGCGAG 2456
QY 2259 GTTCTTCAGTGTGAG 2318
Db 2457 GTTCTTCAGTGTGAG 2516
QY 2319 CTGGAAG 2366
Db 2517 CTGGAAG 2576
QY 2367 ATCAAGCCAGATGAG 2426
Db 2577 ACTGAAGCCGAG 2636
QY 2427 AAGAAAGGAG 2486

Db 2637 AGAAGAGCCAGATGGGAAACAGAAACAGCCCTGTTGAAGCGGAGGCCAACAGGG 2696
 Qy 2487 ATAAATGAGGACGACCTTAATGTCCTCCAGCGCTGCTGCTCTGCTGAGTAAATGCAAGT 2546
 Db 2697 GCCAACGAAGATGACTCTGATGTCCTCGGCGCTGCTGCTCTGCTGATGATGCTGTA 2756
 Qy 2547 GAGAGGAGAAAGATGG-----AAGCCAGGGGAATACGAGAGTGCCTCCAGCTGCTG 2597
 Db 2757 GAAAGGGAGAAAATGGAGGACACAGCAGCCCAAAAAGCCAGAGCAGCCCGAGCAGAAG 2816
 Qy 2598 GGGCTGTGTACGTGTCGAGGAGCTCAGTAAGACTCTGCTCCACACTGTGAGTGTGCA 2657
 Db 2817 GCAGCCACTGAGGTGTCGAGGAGCTCAGCGAGAGTCAAGTTCATATGATGAGCGAGCT 2876
 Qy 2658 GTCAATTGATGGGACCGAGGAGTCAACAGTGTGGAAGCGGCTCTCTCTGCTGGATATCC 2717
 Db 2877 GTCGCTGACGGGACGAGGCGAGCTACCATTTATTGAAGAAGGCTCTCTCTGATATCT 2936
 Qy 2718 GCTTCGTAACAGAACTCTTTGAACACACAGCGGGAGAACCCATGCCACCTGTTGAAGAG 2777
 Db 2937 GCTTCAGTGAAGAACTCTTTGAACAGTGAAGCTGAAGCCGCACTGTTAACTGAGGAG 2996
 Qy 2778 GTCACTGAAGAACACATCTTGC---AGAAGAACTCTCTGCTCACCCAGAGCTTACCA 2834
 Db 2997 GTATTGGAAGAGAAGTAATTGAGAAAGAAACCCCCACGGTTACTGAACCTCTGCCA 3056
 Qy 2835 GAGGCTAAAGATGCCATGACGACATGGTCACCAGTGAAGTGAATTTACCTCAGAAAGCT 2894
 Db 3057 GAGAACAGAGGCCCCGGGGGACACGGTCTGTAGTGAGCGGANTTACCCCCGAAGCT 3116
 Qy 2895 GTGACAGCCACAGAGACTCTCAGAGGCTCTCCGTACTGAAGAGTTACCGAAGCATCGGG 2954
 Db 3117 GTGACAGCTCAGAAACTGACAGGCGCAATTGGGTGCCGAAGAAAGCAACCGAAGCATCTGCT 3176
 Qy 2955 GCCGAGAGACACAGACATGCTGTCGCAAGTTTCCAGCTGACTGACTCCCCAGACACC 3014
 Db 3177 GCTGAAGAGACCAAGAAATGGTGTACGAGCTCTCCAGTTTACCCGACTCCCCAGACACC 3236
 Qy 3015 ACAGAGAGGCCACCCAGTTTCAGGAGGTAGAGTGGTGTGTAGATACAGAAAGAGAG 3074
 Db 3237 ACAGAGAGGCCACTCGGTGTCAGAGGTGGAAGTGGCGTACTGACATAGAAGACAA 3296
 Qy 3075 GAGCGCAGACGAGGCCATCTCTCAAGCGGTTGACAGCAAGGTGAAGAGGAGTCCAG 3134
 Db 3297 GAGAGCGGACTCAAGAGGCTCTCCAGGAGTGCAGCAAAAGTGAAGAGGAATCCAG 3356
 Qy 3135 GTCGCTCAACC---CAGACTGTGACAGAACGGGGTCAAAAGCACTGGAGAGGTTGAG 3191
 Db 3357 CTGCTGGCACCGGTGGCCAGAAATGTTGCTTCAGCTGTGACAGAGCAGAGGCAGAA 3416
 Qy 3192 GAGGTAGAGGAGACTCCGAAGTGTGGCTTCGAGAAAGAGAGGACGTTATGCCGAAA 3251
 Db 3417 AGACCAGAGGACGCTGAAGCTCGGGTCTGAGAAAGAGACGATGATGTTGAAA 3476
 Qy 3252 GGACCGGTGACGAAGCTGAGCTGAGCATCTTGCACAGGGCTCTGAGACTGGACAGGCT 3311
 Db 3477 GTAGATGCTCAGGAGGCAAAACTGAGCCCTTTACACAGGGAAGTGGTGGGCGAGACC 3536
 Qy 3312 ACTCCAGAGAGCCCTTG---AAGTTCCTGAAGTCAAGCAGCATGTA-----GACCAT 3359
 Db 3537 ACCCCAGAAAGCTTTGAAAAGCTCTCAAGTCAAGAGAGCATAGAGTCCAGTGAGCTT 3596
 Qy 3360 GTCGCCACGTGCCA-----GTTTATCAAGCTCCAGCAGCTGATGGAA 3401
 Db 3597 GTAACTACTGTACGCCGAAACCTTAGCTGGGGTAAATATCAAGGAGATGGTATGGAA 3656
 Qy 3402 CAGCGCTGGCCCTGAGTCACTCGAAACCTTGACAGACAGTGAAGCAAAATGGAAGCACT 3461
 Db 3657 CAGGCTATCCCCCTGACTCGGTGGAACCCCTACACACAGTGAAGTGAAGCACC 3716
 Qy 3462 CCCTTAGCAGATTGACACCTGAGATGGGACACAGCAAGATGAACCATTTGACAGCCAG 3521

Db 3717 CCCGTAGCCGACTTTGACGCAACAGGCACAAACCAGAAAGACGAGATTGTGGAAATCCAT 3776
 Qy 3522 GACAGTAAGCCACTGACAGCTGTACGGCAGTCAAGGTCAAGAAAGAGAGCGCGTACT 3581
 Db 3777 GAGGAATGAGGTGCGATCTGGTACCCAGTCAAGGGGGCACAGAACGAGGAGTCTCT 3836
 Qy 3582 GCTCAGAAAGAGGAGGCTTCGACACTACCTAATAATGTTCCAGCCCGAGGAACATGGG 3641
 Db 3837 GCACAGAAAGAGAGGCTCCAGCACCTTCAGTTCAGTTCAGGAAGAAATAAGAA 3896
 Qy 3642 GAAGAACACAGGAAGAGATGTTCTTGAACCTACACAGCAAGAGCTTACTGTCAGCCGTG 3701
 Db 3897 CAATCAAGATGGAAGACACTCTAGAGCATACAGATAAAGAGTGTCACTGGAACATGTA 3956
 Qy 3702 CCCGTTCTCGCAAGACTGAGGTGGGTCAAGAGGGTGA-----GGTTGACTGGTTG 3752
 Db 3957 TCCATTCTGTCAAAGACTGAGGGGACTCAAGAGGCTGACAGATGCTGATGAGAAAC 4016
 Qy 3753 GATGGAAAAAAGTCAAGAAAGAACAGGAGGTGTTGTATACACTCTGG-----ACCCAAC 3806
 Db 4017 AAAGACGTACCATTTTTCGAAGGACTTGAGGGGTCTATAGACACAGGCAATACAGTCAGT 4076
 Qy 3807 AGTCAAAAGGCTGCTGATGTGACATATGAGTAAGTGTGGAGTGCCTGGTCTCAG 3866
 Db 4077 CGGAAAAAGTCACTGAGTTCGCCCTTAAAGTGAAGGACAGAAAGCTGATGTAATAA 4136
 Qy 3867 GAAAAAGAGACTGTAAGTGCAGAG-----TCTTAGCCCTGGAGGAGGGA 3911
 Db 4137 AAGGATGATGCTCTTGAACCTGCAGAGTCAAGTCAAGTCTCTCTCCATCCCGTGGAGAGA 4196
 Qy 3912 GAGATGGAAACTGACCTTGAAGAGGAGAAAGGAGACAAAGCCAGAGCAAGTGAAGTGA 3971
 Db 4197 GAGATGGTGTGTTCAAGTCAAGAGGAGGAGAAACAGAGCAGAGCCCAACCTGTTGAATGA 4256
 Qy 3972 G---AAGGTGACGAGAAACAGCGCTCTGAGCATGAAGGAACCTACGGGAGCCAGTC 4028
 Db 4257 GAGAGCTTGGAGCACAAACAGCTGTTACCGTATCTGAAAGGTCTAGTAAGCAGCTCCT 4316
 Qy 4029 CTGACACTTGAATGCCAGCTCAGAGAGGGGGAAGGCACTGGGAAGCCCTTGGAGGAAGC 4088
 Db 4317 CAGACAGTGAATGTCGCCATCATAGATGGGCAAGGAGTCAAGCAGTTTGGNAGNAGC 4376
 Qy 4089 CTTTCTC---TCCAGACCAAGCAAGCAGGTTGATAGAGTTCAAGTTCAAAGCCCTG 4145
 Db 4377 CTTCTCTCCCTGCTAGGTCAAGAGGAGGAGTATGCACCAAAATTCAGAGTTCAGAGCTCT 4436
 Qy 4146 GACACACAGTCACTCAACACAGCAGAGCTGTGGAAGAGTCTAT-----AGAAACGGTT 4199
 Db 4437 GAGGCATCATTCATCTAAACAGCGCTGCAGAGGAGGAAAGGTCTTAGGAGAAACTGCC 4496
 Qy 4200 GTGATTTACAGACAGGTGAAAAGTCCAGAGTGTGTAGGTGCACACTTATTTACCAGCTGAG 4259
 Db 4497 AACATTTTAGAAACAGGTGAACGTTGGAGCTTCAGAGTGCACATTTAGTTCTGGAAGAG 4556
 Qy 4260 AAGTCTCTGCAACCGGGTGGCCACTGGACTCTTCAGCATGCAAGAGACACGTPACCCCTG 4319
 Db 4557 AAATCTCTGAAAAAATGAAGACTTTGCGGCTCATCCAGGGGAAGATGCTGTGCCACA 4616
 Qy 4320 GGGCTGAGTCTCAGGCAGAAATCCATCCCAATCATAGTAATCTCTCTCTCTGAAAGCACC 4379
 Db 4617 GGGCCGAGTGTGAGCAAAATTCGACACCAAGTATAGTATCTGCTACTACCAAGAAAGGC 4676
 Qy 4380 CTATATCTTGACCTACAAGGAGAAATAAGCGATCCAGAGAGCGGATCAGAGGAAGAG 4439
 Db 4677 TTAAGTTCGACCTGGAAGAGAGAGAAACCATCATCTGAGTGGAGTCAAGTGAAGTC 4736
 Qy 4440 GACAAAGCAGATGCTGGTCTGATGCTGACGGCAAGGAGAGTACAGCAATCGAAAAAGTC 4499
 Db 4737 GATGACAGGTGCTTGGC-----AGGAGTCAAAAGTCAAGTGAAGTGAAGGAGT 4790
 Qy 4500 CTCAGGCTGAACCTGAGATCTCTGGAACCTTGAAGATGAAGCAACAGATTTGTGCTGAAC 4559
 Db 4791 TTAGAGCTGAAAAATGGGATTTTGGAACTTGAGACCAAAAGCAGTAAACTTTGTCCAAAAC 4850


```

Oy 3300 ACTGACAGGCTACTCCAGAGAGCTTG---AAGTTCTGAAGTACGGCAGATGTA--- 3353
Db 3367 GTGGGGAGACCCCGAGAAAGCTTTGAAAAAGCTCTCAAGTACAGAGACATAGAG 3426
Oy 3354 -----GACCATGTGCGCAGCTGCCA-----GGTTATCAAGCTCCAG 3389
Db 3427 TCCAGTAGCTTTGTACACCATTTGTACAGCCGAACCTTAGCTGGGTAAATCAAGAG 3486
Oy 3390 CAGCTGATGAAACAGAGCCGTGGCCCTTGAGTCATCCGAAACCTTGACAGACGTGAGCA 3449
Db 3487 ATGTGATGAAACAGAGCTATCCCTCGACTGTGTGAAACCCCTACAGACAGTGAAGCT 3546
Oy 3450 AATGAGACATCCCTTTGACAGATTTACAGACACTGCAGATGGAGACAGCAAGATGAACC 3509
Db 3547 GATGAGACACCCCGTAGCCAGCTTTGACGACAGGACAAACCAAGAAACAGATTT 3606
Oy 3510 ATTGACAGGACAGATGAAGCCACTGAGCTGTCAAGGACGTCAAGGTCAAGAGAA 3569
Db 3607 GTGGAATTCATGAGAGAGATAGGTCCGATTTGGTACCCAGTCAAGGGGACAGAGAG 3666
Oy 3570 GAGCGGCTACTGTCTCAGAAAGAGAGCCCTTGACACTAATTAATGTTCCAGCCAG 3629
Db 3667 GAGGAGTTTCTGCAAGAGAAAGAGGCTCCAGACCTTCAGATTGTTGTTCCAGGAA 3726
Oy 3630 GAAGAATGAGGAGAAACAGAGAAAGATGTTCTTGAACCTACAGACAGAGCTTACT 3689
Db 3727 GAAACTAAAGAAACAATCAAGATGAGAGACACTTAGACATACATTAAGAGGTCTCA 3786
Oy 3690 GGTGAGCGGTGCGGCTTGCAAAAGACTGAGGTGGTCAAGAGGGGTAG----- 3740
Db 3787 GTGGAATCTGATTCATTTCTGCAAAAGACTGAGGGGACTCAAGAGGCTACAGATAGCT 3846
Oy 3741 GTTGACTGTTGATGAGAGAAAAAGTCAAAAGAAACAGAGAGTGTGTTACACTCTG- 3799
Db 3847 GATGAGAAAAACAAAGACGTACATTTTTCGAAGACTTGAAGGGGTCTATAGACACAGGC 3906
Oy 3800 -----ACCCACAGTCAAAAGGCTGTGATGACATATGACAGTGAAGTATGGAGTG 3854
Db 3907 ATACAGTCAAGTGGGAAAAAGGTCACTGATGATGCTTAAAGGTGAAGGACAGAGAA 3966
Oy 3855 GCGGGGTGCAGAGAAAGAGAGTACTGAAGTGCAGAGT-----CTTACG 3899
Db 3967 GCTGATGTAAAAAGAGATATGCTCTTGAATGACAGAGTCAAGCTTAAGTCTCTCATCC 4026
Oy 3900 CTGAGAGAGAGAGATGAGAACTGACGTTGAAAAAGAGAAAAAGGAGCAAAAGCCAGAG 3959
Db 4027 CCGGTGAGAGAGAGATGATGTTCAAGTCAAGAGGAGAAAAAGAGACAGAGCCAAACC 4086
Oy 3960 CAAGTGAAGTAAAG---AAGGTGACAGAGAAAGCCGCTCTCTGAGCATGAAGAACTTAC 4016
Db 4087 CATGTGAATGAGAGAGAGCTTGAACGAGAAACAGCTGTATCCGTTATCTTAAGAGGTCACT 4146
Oy 4017 GGGAAAGCCAGTCTTCACTTGAACATGCCCAGCTCAGAGGGGAGAGGACCTGGGAAGC 4076
Db 4147 AAGCAGCTCTCCAGACATGATGATGCCCCATATGATGAGGGCAAGAGAGTCAAGAT 4206
Oy 4077 CTGAGAGAGAGCCCTTCTC---TCCAGACCAAGCAAAAGAGTTCATAGAGTTCAA 4133
Db 4207 TTGGAAGAGAGCCCTCTCTCCCTGCTAGTCAAGAGAGAGAGATGACCAAAATTCAA 4266
Oy 4134 GTTCAAAAGCTTGACCAACAGTCACTCAAAACAGAGAAAGCTGTGAAAAAGTCAAT--- 4189
Db 4267 GTTTCAGAGCTCTGAGGATCATCTTCAACAGCGCGCTGACAGAGAGAGAAAAAGGCTTCA 4326
Oy 4190 --AGAAACGTTGTGATTTCAAGACAGGTGAAGTCCAGAGTGTAGGTGACACTTA 4247
Db 4327 GGAAGAAATGCAACATTTTGAAGAGGTGAAGAGCTTGAAGCTCAGAGTGCACATTTA 4386
Oy 4248 TTACAGCTGAGAGATCTCTGCAAGCGGTGAGCACTGAGACTTTTCAGAGTCAAGAGAC 4307
Db 4387 GTTTCGAGAGAGAAATCTCTGAAAAAAATGAAGACTTTTCCGCTCATCAGAGGAGAAAT 4446

```

```

Oy 4308 ACGTACCCCTGGGGCTGAGTCTCAGAGAAATCCATCCCAATAGTAATCTCTGCT 4367
Db 4447 GCTTGCCCAACAGGCGCCCACTGTCAAGCAAAATCCAGACAGATAGATATCTGCTACT 4506
Oy 4368 CTTGAAAGCACTTACATCTGACCTTACAGAGAAATTAAGCGATCCAGAGAGAGCA 4427
Db 4507 ACCAAGAAAGCTTAAAGTCCGACCTGGAAGAGAGAAAAACATCATCAAGATGAGAG 4566
Oy 4428 TCAGAGAGAGAGCAAGCCAGATGCTGTCCTGATCTCTACACGAGAGAGATGACGA 4487
Db 4567 TCAGATGAAGTCAAGTGAAGAGGTTGCTTGC-----AGAGGTCAAGATGATGAGCA 4620
Oy 4488 ATCAAAAAAGTCTCAAGGCTGAACCTGAGATCTGGAACCTTGAGATTAAGAGCAAG 4547
Db 4621 ATTAGAGAGATTTAGAGCTTGAAAAATGGATTTTGAACCTTGAGACCAAAAGCATGAA 4680
Oy 4548 ATTGTGCTGAACGTATTCAGACAGAGCCGTTGACAGTTGCAAGCTAC--AGAAACAGCC 4604
Db 4681 CTTGTCCAAAACATCATCTCAGACAGAGCCGTTGACAGATTGTAAGTACAGAGAAACAGCC 4740
Oy 4605 CCGGAATCTAGCTTATGATTCACAGACCCAGGTTCTCTGATCAGAGCTTGAACAGAG 4664
Db 4741 ACCGAAATGTTGAGTCTGAGTTACAGACCAAGCTCAGCTGATTAAGAGTGAAGCAG 4800
Oy 4665 GA 4666
Db 4801 GA 4802

```

```

RESULT 6
US-09-724-676A-12736
; Sequence 12736, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: Patent version 3.2
; SEQ ID NO 12736
; LENGTH: 5773
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (91)..(91)
; OTHER INFORMATION: n is a,c,g, or t
US-09-724-676A-12736

Query Match      31.2%; Score 1920.8; DB 5; Length 5773;
Best Local Similarity 68.2%; Pred. No. 0;
Matches 3072; Conservative 0; Mismatches 1247; Indels 183; Gaps 22;

Oy 227 GTTGACAGCGAGAGTCAAGATGTGAGAAAAAGACGAGTTGAAGAAATGCGGCGC 356
Db 332 GTTGACAGAGAGTCTGAGAGATGTGAGCAAGAGACTCCGATTAAGAGATGCTACT 411
Oy 357 AACTCCAGAGCTGTTGAGATATCAAAAGAGATGGCGAGAGAGACATCAAGAAATAAT 416
Db 412 AAGTCAAGGGTGTTCACAGACATCAAGATGATGGCGAGAGAGACCCGAAATAATTC 471
Oy 417 GAACGATCCCTGCTTCAGAAAACAATGTGAGAAATGTGACAGCTGTGATGCCAG 476
Db 472 GAACGATTCCTTCTTCAGAAAGCAATTTAGAGAGCTTACACAAACCATGAGTCCAG 531
Oy 477 GCTATGATGTGGCTTCAAGAAAGATTTAAATTTGTGTTTAAATTCACGGTGAAG 536
Db 532 GCTAATGATATTGATTTAAAGAGGTGTTTAAGTTGTGGCTTTAAATTCATGTGAAA 591
Oy 537 AAGATTAAGAAATGAAGATCAATCTGTCCAACTTCACTGTCTCAAGAGATGAAGGC 596
Db 592 AAGATTAAGACAGAGAACCTGTGACATGTCCAGGTACTCACTGTGAAAGAAAGATGAAGG 651

```

Qy	597	GAAGGGCAGAACGCTCTGTCTGAGAGTGTGAGACACACAGGAGGCCAGTGTGTGGAGACTGCC	656
Db	652	GAGGAGCAGCA-----GGGGCTGGCGACCCACAGGACCCACAGCTTGGGGCT---	699
Qy	657	GTCCGAGAGTTCAGATCCAAAGAAAGTGAGCTGTGAGCAATCCACAGAGAAGCAAGAGGC	716
Db	700	---GGAGAGCAGATCCAAAGAAAGCAACCAACAACTCTACAGAGAAACCCGAGAG	756
Qy	717	ACCTGAAGCAAGAACAGAGCAGCAGAGAAATCCCCCTTCAAGCCGAATCTGTATCAAGCG	776
Db	757	ACCCTGAAGCGTGAGCAAGCCAGCAGAAATTTCTCCCCAGCGCAATCTGGCCAA---	813
Qy	777	GCTGAGAGAGCCAAAGATGAAGAGAGAGAAACAAAGAGAAAGAGCCCAACAGTCC	836
Db	814	GCAGTGGAGGAATGCAAGAGGAGAGAGAGAGAAACAGAGAAAGAAACCTTAGCAAGTCT	873
Qy	837	CCAGAAATCCCGAGAGCCCGAGTCAACGTGTGAGACAAATCTTCTCTCAAGAGTCTCTC	896
Db	874	GCAGAAATCTCCGACTAGTCCCGTGACCGTGTGAACAGAGATCAACCTTCAAAAATTTCTTC	933
Qy	897	ACTCACGGTTGGGCGGCTGGCCAAAGAGACACAGCTTCAAGAAATCAAAAGAGAGATGAT	956
Db	934	ACTCAAGTTGGGCGGCTGGCGCAAAAGACCAAGTTTTCAGGAGCCGAGAGAGATGAA	993
Qy	957	CTGAAATCTCCGAGAGAGAAAGGAGCAGAGGCGAGAGAAAGTAGACAGAGAAAGAAAG	1016
Db	994	GTGGAAGCTTCAGAGAGAGAAAGGAAACAGAGCCAGAGAAAGTAGACAGAGAGAGAC	1053
Qy	1017	GAAGAAGCAGAGCCAGCTCGAGGA-----CGAGAG	1049
Db	1054	GGAAAGCGAGGTTGCTCCGAGAAACTGACCGCTCCGAGCAAGCCACCCACAGGAG	1113
Qy	1050	CCGGCAGAGACACAGACACAGGCCAGGTCAGTGTGTGACAGACTACGAGAAGGTGAGCTGCCT	1109
Db	1114	CCGGCAGAAAGTGCACAGGCCCGGTTATCAGCTGAATATGAGAAATTTGAGCTGCC	1173
Qy	1110	TTGGAAGACCAAGTTGGTGACCTGGAGGCATCGTCAGAGGAGAGTGTGTCTCTTGGCA	1169
Db	1174	TCAGAGGAGCAAGTCAGTGGCTCGCAGGACCTTCTGAAGAGAAACCTGCTCCGTGGCG	1233
Qy	1170	ACGGAAGTGTTCATGAGAAGATGGAAGCCCAACAGAA---GTTGTGTGAGAGGTCCAC	1226
Db	1234	ACAGAAGTGTTCATGAGAAATATAGAAGTCCCAACAGAGAGGTTGTGGCCGAAGTCCAC	1293
Qy	1227	GTGAGCACCGTGGAGAGAGACAGAGGAGGAGCAGGAGGAGGAGAGGCTGAAGGGGGC	1286
Db	1294	GTGAGCACCGTGGAGGAGAGACCGAAGAGCAAA-----	1329
Qy	1287	GTGTGTGTAGAAGGAACAGGAGAAATCTTTGCCCCCTGAGAAACTGTGGCTGAGCCCCAGGAG	1346
Db	1330	ACGGAGGTGAAGAAACAGCAGGGTCTGTGCCAGCTGAAGAAATTTGTTGAAATGGATGCA	1389
Qy	1347	GTCCCCCAGGAAGCTGAGCCTGTGAGGAGCTGTATGAGAGCAGAGAGATGTGTCTCT	1406
Db	1390	GAACCTCAGGAAGCCGAACCTGCAAGGAGCTGGTGAAGCTCAAGAAACGTTGTGTTC	1449
Qy	1407	GGAGGAGACCACATCACTGACAGACCTTAAGTCTGTGAGAGAGAGACGCTGCCCAACAC	1466
Db	1450	GGAGGAGACCTACACAGGAGCTGACCTCAGTCTGTATGAGAAGGTGCTGTCAAAACCC	1509
Qy	1467	CCAGAAGGCATTTGTCAGTGAGGTGGAGATGCTGTCTCTTCAGGAAAGAAATCAAGGTACAG	1526
Db	1510	CCGAGGCGTTGTGAGTGAGGTGGAATATGCTGTATACAGGAGAGAAATGAGTGCAG	1569
Qy	1527	GGAAAGTCCCTTGAAGAAATCTTTTCAAGTGTGAGTGTGAGTGTGAGGAGAGAG	1586
Db	1570	GGAAAGTCCAATAAGAGCTTTTACAGCACTGGCTTAAAGAAAGCTTTCTGGAAGAAA	1629
Qy	1587	CAGAAAGGGGAAACGAGGAGTGGGGAGACGAAGAGCCTGTGAGAAATACCAACATTTAC	1646
Db	1630	CAGAAAGGGGAAAGAGGA---GGAGGAGACGAGGAATCAGGGGAGCACATCTCAGTTC	1686

QY	1647	ACCGAATCCCAGAGAGTGTGATGACAGAAAGGAGAGAGGCTCTGCGTCTGCCCGAG	1706
DB	1687	GCCGATTTCTCCGACAGCCAGGAGGAGCAAAAAGGCGAGAGGCTCTGCGCTCATCCCGCTGAG	1746
QY	1707	GAGCCTCAGGAGACACGTTGTCTGGAGAAGAGGCGCGCTGGGAAGCACCCACGAGTGGGAA	1766
DB	1747	GAGCCCGAGAGATCAGCTGTCTGGAAAAGGCGTTAGCCGAGGTGCACAGGNTGGGAA	1806
QY	1767	GCTCAGGAAGAACTACTTCCGATGGAGAGAAGAAGAGGAGTCACTCCCTGGGCA	1826
DB	1807	GCTGAAGAAGAGCTACTTCCGATGGAGAGAAAAAAGAGAAGGTGTCACTCCCTGGGCA	1866
QY	1827	TCCTTCAAAAAGATGTGACACCCCAAGAAAACGGGTCGGAAGACCTTCTGAGAGTGCACAG	1886
DB	1867	TCATTTCAAAAAGATGTGTGACGCCCAAGAAGCGGTGTAGACGGCCTTCGGAAGGTGATAAA	1926
QY	1887	GAGCAAGAGCTGGAGAAGGTCAAGAGCGCCACCTTTGTCTCTCCACTGTATGACACAGTGTCA	1946
DB	1927	GAGATGAGCTGGACAAAGGTCAAGAGCGCTACTTTGTCTTCCACCGAGAGCACAGCCTCT	1986
QY	1947	GAATTCGAAGATGAAGTCAAAACTTTGTGTAGGAACAAAAAGCCAGAGGAACCAAGACGT	2006
DB	1987	GAATTCGAAGAAATGAAGAGGCGGTGAAGAGGCCAAAGCCGGAAGAACCAAGACGC	2046
QY	2007	AGGTTGATACTTACGTGTCTTGGGAAGCACTGATTTGTGTGCGATCATCCAGAAGAGA	2066
DB	2047	AAGTTGATATCTCAGTATCTTTGGGAAGCTTTAAATTTGTGTGGATCATCCAGAAGAAAGA	2106
QY	2067	GCAAGGAAGCATCTCTTCAGATGATGAAGAGGGGCCAAGGACACTGGGAGGGGACAGT	2126
DB	2107	GCAAGGAGAGGGTCTCTTCTGATGAGGAAGGGGACCAAAAGCAATGGGAGGAGACCAAC	2166
QY	2127	CACAGACAGAGGAGGCCAGCAAGAACAAGAAGAGCCGGAACAGACGCTGTTCTCTGCCAGC	2186
DB	2167	CAGAAAGCTGATGAGCCGGAAGAACAAAGAGACGGGGAACAGCGGATCTTGTCTGCT	2226
QY	2187	ACCCAGAGCAGGACCAAGCGCAAGGAAGTCTTCAACCGAGCCAGCGGAAGCCCTTCC	2246
DB	2227	TCCCAAGAACATGATCAGGGCAGGGAAGTCTCTCCCGAGCAAGCTGGAAGCCCTACC	2286
QY	2247	GAAGGGAAGGTCTCCACTTGGGAGTCAATTTAAAGNTAGTCACTCCGAAGAAAAA	2306
DB	2287	GAAGGGAGGGCGTTTCCACTCTGGGAGTCAATTTAAAGGTTAGTCAACGCCAAGAAAAA	2346
QY	2307	TCCAAGTCAAAACTGGAAGAGAAGCCGAAGAC-----TCTAGTGTAGAGCAG	2354
DB	2347	TCAAAGTCCAAGCTGAGNAGAAAACGGAAGACTCCATAGCTGGGTCTGGTGTAGACAT	2406
QY	2355	TTGTCCACTGAGATTCGAACCGAGTAGAGAAGAAATCTTTGGGTTTCATTTAAGAAATTCATC	2414
DB	2407	TCCACTCCAGACACTGAACCCCGTAAAGAAGAAATCTCTGGGTCTCAATCAAGAAGTTTATT	2466
QY	2415	CCCCGAGCGGAGAAAGAGGCGACACGGGAGCAGACAGCAAGCCACTGTCTGGAAGACTCA	2474
DB	2467	CCTCGACGAAGGAAGAAAAGCCAGATGGGAACCAAGAACCAAGCCCTCTTTGAAGAAGCGCA	2526
QY	2475	GGGCGAGTGGAGATAATGAGGACGACCCCTAATGTCCAGCGGTCTGTCCTCTGTCTGTGAG	2534
DB	2527	GGGCAACACAGGGCCACAGNATGACTCTGATGTCTCCGCGGTGTCTCTGTCTGTGAG	2586
QY	2535	TATAATCAGTGGAGGGGAGAAGATGG-----AAGCCCCAGGGGAATACGAGACTG	2585
DB	2587	TATGATGCTGTAGAAAAGGAGAAATGGAGGCACAGCAAGCCCCAAAAAAGCGCAGAGCAG	2646
QY	2586	CCCAGCTGTCTGGGGCTGTGTACGTGTCCGAGGAGCTCAGTAAGACTCTGTGGTCCACACT	2645
DB	2647	CCGAGCAGAAGGACGCCACTGAGGTGTCTCAAGGAGCTCAGCGAGTCAAGTTCATATG	2706
QY	2646	GTGAGTGTCTCAGTTCATTGATGGGACCGAGGAGTCAACAGTGTGGAAGAGCGGTCTCTCT	2705
DB	2707	ATGCGCAGACTGTCTGCTGACGGGACGAGGGCAGTCACTTATTTAAGAAAGGTCTCTCT	2766
QY	2706	TCGTGGATATCCGCTTCCGTTAACAGAACTCTTTGAACACACAGCGGGGAGAACCATGCCA	2765

Db 2767 TCTTGGATATCTGCTTCAGTACAGAACCTCTTGAACAAATGTAAGACCTGAAAGCCCACTG 2826
 Qy 2766 CCTGTGAAGAGTCACTGAAAAAGACATCATTTGCA---GAAGAACTCTGTGTCAAC 2822
 Db 2827 TTTAACTGAGAGATTTTGGAAAAGAAATTAATTGCAGAAAGAAAGAACCCCAAGTTTACT 2886
 Qy 2823 CAGAGCTTACAGAGAGGTTAAAGATGCCATGACGACATGTGTCAACAGTGAAGTGAATTTT 2882
 Db 2887 GAACCTCTGCAGAGAAACAGAGAGGCCCGGCGGCGACACGCTGTGATGAGGCGGAATTTG 2946
 Qy 2883 ACCCTGAGAGCTGTGACAGCCACAGAGCCTGAGAGGCTCCCTGACTGAAAGATTAAC 2942
 Db 2947 ACCCTCGAAGCTGTGACAGCTGCAAGAACTGCAAGGCCATTGGGTGCCGAAAGAGAAAC 3006
 Qy 2943 GAAGCATCGGGGCGGAAAGAGACACAGACATGTGTCCGAGTTTCCAGCTGACTGAC 3002
 Db 3007 GAAGCATCTGCTGTGAAAGAGACACAGAAATGTGTCAAGAGTCTCCAGTTAACCGAC 3066
 Qy 3003 TCCCCAGACACCAAGAGAAAGCCACCCCAAGTTCAAGAGTGAAGAGTGTGTCTAGAT 3062
 Db 3067 TCCCCAGACACCAAGAGAGGCCACTCCGGTGACAGAGTGAAGTGGCGTACCTGAC 3126
 Qy 3063 ACAGAAAGAGAGAGAGCCGACAGCGAGGCATCTCCAAAGCCGTTGACAGCAAGGTGAA 3122
 Db 3127 ATGAAAGAGCAAGAGAGGCGGACTCAAGAGGTCTTCAGGCAGTGGCAGAAAAAGTTGAA 3186
 Qy 3123 GAGAGTCCCAAGTCTCTGCAACCA---GACTGTGACAGAAACGGGGTCAAAAGCACTG 3179
 Db 3187 GAGGAATCCCACTCTCTGTGACACCGGTGGCCAGAAAGATGTCTTCAAGCTGTGTGACAGA 3246
 Qy 3180 GAGAAGTTGAGAGAGTGAAGAGAGACCTCCGAAGTGTGGCTTTCGAGAAAGAGAGAC 3239
 Db 3247 GAGAGGCGCAAGAAAGACCAAGAGAGCGTGAAGCGTGGGTCTGAAGAAAGACGAGAT 3306
 Qy 3240 GTTATGCCGAAGAGACCCGTGAGAGAGTGAAGTGAAGCATCTTGACAGGGCTCTGAG 3299
 Db 3307 GTAGGTGTGAAGTGAATGATGCTCAGAGAGCAAAAGCTGTAACCTTTTACACAGGAAAGT 3366
 Qy 3300 ACTGACAGGCTACTCTCAGAGAGCCTTG---AAGTTCTGAAGTCAACGACAGATGA--- 3353
 Db 3367 GTGGGGCAGACCAACCCAGAAAGCTTTGAAAAAGCTCTCAAGTCAACAGAGCATAGAG 3426
 Qy 3354 -----GACCATGTGCCACGTGCCA-----GTTATCAAGCTCCAG 3389
 Db 3427 TCCAGTAGCTTGTAAACCACTGTCAAGCCGAAACCTTAGCTGGGGTAAATCAACAGAG 3486
 Qy 3390 CAGCTGATGGAACAGGCCGTGGCCCTGTGATCCGAAACCTTGACAGACAGTGAACA 3449
 Db 3487 ATGTGTATGGAACAGGCTATCCCTCTGACTGTGTGAAACCTTACAGACAGTGAAGCT 3546
 Qy 3450 AATGGAAGCACTCCCTTACAGATTCAGACATCTGACATGAGGACACAGCAAGATGAACC 3509
 Db 3547 GATGAAGACACCCCGTAGCCGACTTTGACGACAGGACCAACCCAGAAACAGAGATTT 3606
 Qy 3510 ATTGAACGCCAGAGACATGAAGCCACTGACGTGTCAAGGACGTCAACAGTCAACAGAA 3569
 Db 3607 GTGGAATATCATGAGAGAAATGAGTGCATCTGTAACCAAGTCAAGGGGACACAGAA 3666
 Qy 3570 GAGGGGCTACTGTCTCAGAAAGAGAGCTTGAACATCACTAATTAATTTTCAAGCCAG 3629
 Db 3667 GAGGCAATCTCTGACACAGAAAGAGGCTCCAGCACTTCCAGTTTGTGTTCCAGGAA 3726
 Qy 3630 GAAGAACATGGGGAGAGAACACAGAAAGAGATGTTCTTGAACCTACACAGCAAGAGTTACT 3689
 Db 3727 GAAACTAAAGAACATCAAGATGAGAGCACTTAGACATACAGATTAAGAGGTGTCA 3786
 Qy 3690 GCTGACGCCGTGCCGTTTTCGCAAAAGTGAAGTGAAGGTCAAGAGGTGAG----- 3740
 Db 3787 GTGGAACACTGTATTCATTTCTGTCAAAAGCTGAGGGGACTCAAGAGGCTGACCACTATGCT 3846
 Qy 3741 GTTGAAGTGTGATGAGAAAAAGTCAAAAGAGAAACAGAGAGTGTGTTGTAACACTGTG- 3799

Db 3847 GATGAGAAAACAAAGACGTACCACTTTTTCGAAAGACTTGAAGGGGTCTATAGACACAGGC 3906
 Qy 3800 -----ACCCAACGTCAAAAGCTGTGATGTGACATATATGACGTGAAGTGAAGGTG 3854
 Db 3907 ATAACTGACGTGGGAAAAAGTCACTGAAGTTGCCCTTAAAGTGAAGGACAGAAAGAA 3966
 Qy 3855 GCCGGTGTCAAGGAAAAAGAGAGTGAAGTGCAGAGT-----CTTAGC 3899
 Db 3967 GCTGAATGTAAAAAGATGATGCTCTTGAACCTGACAGTCAACGCTAAGTCTCTTCAATCC 4026
 Qy 3900 CTGAGAGAGGAGAGATGAGAACTGACGTTGAAAGAGAGAAAAAGGAGAACCAAGCCAG 3959
 Db 4027 CCGTGGAGAGAGATGTGATTTCAAGTGGAAAGGAGAAAAAGAGAGACAGAGCCAAAC 4086
 Qy 3960 CAATGTGTGAAG---AAGTGAACAGAAACAGCCGCTCTTGAGCATGAAAGAACTTAC 4016
 Db 4087 CATGTGAATGAAGAAAGCTTGAGACAGAAACAGCTGTACCGTATCTGAAGAGTGAAGT 4146
 Qy 4017 GGAAGCCAGTCTGACATGACCTTGAACGCTCCAGAGGAGGGGAAAGGACCTGGAAAGC 4076
 Db 4147 AAGCAGCTCTTCAAGACATGAATGTGCCATCATATAGATGGGCGCAAGAGAACTGACAGT 4206
 Qy 4077 CTGAGAGAAAGCCCTTCTC---TCCAGACCAAGACAAAGCAGGTTGCATAGAGTTCAA 4133
 Db 4207 TTGGAAGGAAAGCCCTCTCCCTGCTTAGGTCAAGAGAGGAGATGATCCAAATTTCAA 4266
 Qy 4134 GTTCAAAAGCTGACACACACATCTCAACACAGAGAGCTGTGAAAAAGTCAAT--- 4189
 Db 4267 GTTCAGAGCTGTGAGCATCATCTCACTTAAACAGCGGCTGCAGAGGAGAAAGTCTTA 4326
 Qy 4190 --AAGAAAGCTGTGATTTCAAGACAGGTGAAGTCCAGATGTGTAGGTGACACCTTA 4247
 Db 4327 GGAAGAACTGCAACATTTTGAAGAAAGGTGAACGTTGAGCTGCAAGGTGACATTTA 4386
 Qy 4248 TTACAGCTGAAGAGTCTCTGCAACCGGTGGCCACTGACCTTTCAGCATGTGACAGAGAC 4307
 Db 4387 GTTCTGAGAGAGAAATCTCTGAAAAAATGAAGACTTTGCCGTATCCAGGGAAAGAT 4446
 Qy 4308 ACGGTACCCCTGGGGCTGAGTCTCAGGCAAGATCCATCCCATCATATGTAATCTCTGCT 4367
 Db 4447 GCTGTGCCCAAGGGCCCGACTGTCAAGCAAAATCGACACAGATGATGATCTGTACT 4506
 Qy 4368 CCTGAAAGCAACCTTACATCTGACCTTCAAGGAGAAATTAAGCGCATCCAGAGAGAGCA 4427
 Db 4507 ACCAAGAAAGGCTTAAATTTCCGACCTGAGAGAGAGAAACCATATCATTAAGTGAAG 4566
 Db 4567 TCAATGTAAGTGAATGAGACAGTGTGCTTGGC-----AGAGGTCAAAAGTGAAGTGA 4620
 Qy 4488 ATCGAAAAAGTCTCAAGAGCTGAACCTGAGATCTCGAATCTTGAAGTGAAGCAACAG 4547
 Db 4621 AATGAGAGAGATTTAGAGCTGAAATGGAATTTTGAACCTTGAAGCCAAAGAGTAAA 4680
 Qy 4548 AATGTGCTGAAGTCAATTCAGACGCGTGTGACAGTTCGACGATC---AGAAACAGCC 4664
 Db 4681 CTGTGCAAAACATCATCAACAGCGGTTGACAGATTTGATGACAGAAAGAGCC 4740
 Qy 4605 CCGCAAACTCATGTATGATTTCAACAGCCAGAGTTCCTGCATCAGGCTTGAACAGCAG 4664
 Db 4741 ACCGAAATGTTGACGTCTGAATTTACAGACACAACTCAACGTGAATTAAGTGAAGCAG 4800
 Qy 4665 GA 4666
 Db 4801 GA 4802

RESULT 7
 US-10-152-319A-1312/c
 ; Sequence 1312, Application US/10152319A
 ; GENERAL INFORMATION:
 ; APPLICANT: Mendrick, Donna
 ; APPLICANT: Porter, Mark

; APPLICANT: Johnson, Kory
 ; APPLICANT: Higgs, Brandon
 ; APPLICANT: Castle, Arthur
 ; APPLICANT: Elashoff, Michael
 ; TITLE OF INVENTION: Molecular Toxicology Modeling
 ; FILE REFERENCE: 44921-5089-US
 ; CURRENT APPLICATION NUMBER: US/10/152,319A
 ; PRIOR FILING DATE: 2002-05-22
 ; PRIOR APPLICATION NUMBER: US 60/292,335
 ; PRIOR FILING DATE: 2001-05-22
 ; PRIOR APPLICATION NUMBER: US 60/297,523
 ; PRIOR FILING DATE: 2001-06-13
 ; PRIOR APPLICATION NUMBER: US 60/298,925
 ; PRIOR FILING DATE: 2001-06-19
 ; PRIOR APPLICATION NUMBER: US 60/303,810
 ; PRIOR FILING DATE: 2001-07-10
 ; PRIOR APPLICATION NUMBER: US 60/303,807
 ; PRIOR FILING DATE: 2001-07-10
 ; PRIOR APPLICATION NUMBER: US 60/303,808
 ; PRIOR FILING DATE: 2001-07-10
 ; PRIOR APPLICATION NUMBER: US 60/315,047
 ; PRIOR FILING DATE: 2001-08-28
 ; PRIOR APPLICATION NUMBER: US 60/324,928
 ; PRIOR FILING DATE: 2001-08-28
 ; PRIOR APPLICATION NUMBER: US 60/330,867
 ; PRIOR FILING DATE: 2001-11-01
 ; PRIOR APPLICATION NUMBER: US 60/330,462
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 2221
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1312
 ; LENGTH: 563
 ; TYPE: DNA
 ; ORGANISM: Rattus norvegicus
 ; FEATURE:
 ; OTHER INFORMATION: Genbank Accession No. A1233818
 US-10-152-319A-1312

 Query Match 8.9%; Score 550; DB 6; Length 563;
 Best Local Similarity 99.1%; Pred. No. 3.1e-120; Indels 0; Gaps 0;
 Matches 553; Conservative 0; Mismatches 5;

 QY 5601 ATACGACGCTTTGAAAAGCTCCAGCCCTCAACTGTAACCTGCAGCAACACAGATAACATT 5660
 DB 563 AAAGCGACGCTTTGAAAAGCTCCAGCTTCAACTGTAACCTGCAGCAACACAAATCAATT 504
 QY 5661 CCTGCAAGACAGACAAGTCTTTTTTAAAGTTTAACTGATGCTTATGATCTGTGGCTTCT 5720
 DB 503 CCTGCAAGAAAAGACAAGTCTTTTTTAAAGTTTAACTGATGCTTATGATCTGTGGCTTCT 444
 QY 5721 AGTCCTCTGAAAGTGGTTGTTTCTCTATGCACAGCGAGCTCAGAAATAAACCACCATTT 5780
 DB 443 AGTCCTCTGAAAGTGGTTGTTTCTCTATGCACAGCGAGCTCAGAAATAAACCACCATTT 384
 QY 5781 TGAACATCCAGGATGTCCCAATATTACCATGATTTTTTCCCCCTCTTTTGTCTAATCCAG 5840
 DB 383 TGAACATCCAGGATGTCCCAATATTACCATGATTTTTTCCCCCTCTTTTGTCTAATCCAG 324
 QY 5841 TCCAGGTTGGAAGAAGTCTCTCTGTGTGATTAAGCCCTGTCTCTTAATGATATGA 5900
 DB 323 TCCAGGTTGGAAGAAGTCTCTCTGTGTGATTAAGCCCTGTCTCTTAATGATATGA 264
 QY 5901 CAATAGTGTGCTTAAGCCCATGAGATGTTTCTTAATGCAGGAAGATCTGTCTAGT 5960
 DB 263 CAATAGTGTGCTTAAGCCCATGAGATGTTTCTTAATGCAGGAAGATCTGTCTAGT 204
 QY 5961 TTTTGTGATGCTCTCTCTATGCTGACCGCAATTCATATGCAGATCGAAGTGTCTGT 6020
 DB 203 TTTTGTGATGCTCTCTCTATGCTGACCGCAATTCATATGCAGATCGAAGTGTCTGT 144
 QY 6021 TCTTTACAGATGGTATTTTGTGATGATCTGGAGTTTGTCTGTGTATATCTGTGCCCTT 6080

```

RESULT 9
US-10-152-319A-1193/c
; Sequence 1193, Application US/10152319A
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Potter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Higgs, Brandon
; APPLICANT: Castle, Arthur
; APPLICANT: Elashoff, Michael
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5089-US
; CURRENT APPLICATION NUMBER: US/10/152,319A
; PRIOR FILING DATE: 2002-05-22
; PRIOR APPLICATION NUMBER: US 60/292,335
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/297,523
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,925
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,810
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/303,807
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/303,808
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/315,047
; PRIOR FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: US 60/324,928
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/330,867
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: US 60/330,462
; PRIOR FILING DATE: 2001-10-22
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2221
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1193
; LENGTH: 180
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. AI228494
US-10-152-319A-1193

Query Match          2.0%; Score 124.6; DB 6; Length 180;
Best Local Similarity 83.0%; Pred. No. 6.4e-20;
Matches 142; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 5987 GACCGAATTCATATGCAATGCAAGTACCTGTTCTTTACAGATGTAATTTGATAGA 6046
DB 180 GGGCGAATTAATTAATGCAATGCAAGTACCTGTTCTTTACAGAAAGGATTTGAAAAA 121
QY 6047 TACTGAGATTTGTCTGTATATCTGTGTCCTTCTTTAAACAATGTCATTATGT 6106
DB 120 AACTGGGGTTGGTGGTTAAATTTGGGCCCCCTTTTAAAAACAATTTGCTTAAGT 61
QY 6107 TCCTTGGATAATGTGATTTGACACACGATTTAAATTAACAATTTTAC 6157
DB 60 TCCCTTGAATAATTTGGATTTGCCACTGATTTTAAATAACAATTTGCC 10

RESULT 10
US-09-724-676-12734
; Sequence 12734, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2

```

```

; SEQ ID NO 12734
; LENGTH: 753
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676-12734

Query Match          1.9%; Score 119; DB 5; Length 753;
Best Local Similarity 64.7%; Pred. No. 2.4e-18;
Matches 214; Conservative 0; Mismatches 105; Indels 12; Gaps 2;

QY 3 GGGCGGGGAGTAGAAGACCACTGAGCCATGCGCGCAGAGATTTCACAGACAGCG 62
DB 162 GGCTAGGCGGGAGAAAGTCCGAGAGCCATGGGCGCGGAGCTTCACCGAGACGCG 221
QY 63 AGCCCGAGACGCTG---GGGGGAGGCAACGCCAGCGAGCTGTCTCAGTGGCCAT 119
DB 222 AGCCCGAGAGACGCGCGCGGAGGAGCTCCACGCGGCTAGCCCGAGCGGCGGC 281
QY 120 GGGCGCGAGCTGAAGCTCGGGA-----GCAGCTGAGACCCCGCGAGCGGAGC 170
DB 282 GGGCGCTGCGCGGAGGCGGCGCCAGACACACCGGAGACCCCGCATGCTGCTCGGAC 341
QY 171 CCCGCAACCAAGCTCCCAAGAAAGATGCGCAGCTGTTCTGTCAACGGCGTAGTGA 230
DB 342 CCCGCAACCAAGCTCCTAAGAAAGATGTCAGCTGTCCACATCAATGGGTAGCTGAG 401
QY 231 CAAGAGATGTCATGTCACAGAGAAAACGAGAGGGGCGAGGAGAAAGTCTTGAT 290
DB 402 CAAGATGAGCTCAAGCTTCAGAGAGGTGACCTTAATGGCCAGAAAGAGCCCTGAC 461
QY 291 GAGGATGTTGACAGCGAGAGTCAAGATG 321
DB 462 CAAGAGAGCTTAACAGCGAGAGAAAGAG 492

RESULT 11
US-09-724-676A-12734
; Sequence 12734, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12734
; LENGTH: 753
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676A-12734

Query Match          1.9%; Score 119; DB 5; Length 753;
Best Local Similarity 64.7%; Pred. No. 2.4e-18;
Matches 214; Conservative 0; Mismatches 105; Indels 12; Gaps 2;

QY 3 GGGCGGGGAGTAGAAGACCACTGAGCCATGCGCGCAGAGATTTCACAGAGAGCG 62
DB 162 GGCTAGGCGGGAGAAAGTCCGAGAGCCATGGGCGCGGAGCTTCACCGAGAGCGC 221
QY 63 AGCCCGAGACGCTG---GGGGGAGGCAACGCCAGCGAGCTGTCTCAGTGGCCAT 119
DB 222 AGCCCGAGAGACGCGCGCGGAGGAGCTCCACGCGGCTAGCCCGAGCGGCGGC 281
QY 120 GGGCGCGAGCTGAAGCTCGGGA-----GCAGCTGAGACCCCGCGAGCGGAGC 170
DB 282 GGGCGCTGCGCGGAGGCGGCGCCAGACACACCGGAGACCCCGCATGCTGCTCGGAC 341
QY 171 CCCGCAACCAAGCTCCCAAGAAAGATGCGCAGCTGTTCTGTCAACGGCGTAGTGA 230
DB 342 CCCGCAACCAAGCTCCTAAGAAAGATGTCAGCTGTCCACATCAATGGGTAGCTGAG 401
QY 231 CAAGAGATGTCATGTCACAGAGAAAACGAGAGGGGCGAGGAGAAAGTCTTGAT 290

```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 13, 2002, 00:12:21 ; Search time 329 Seconds
(without alignments)
3127.641 Million cell updates/sec

Title: US-09-902-432-4
Perfect score: 8073
Sequence: 1 MNGSSTEQSPQAGSDT.....AWAQRKCLPRLQKAPVSK 1596

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA Main.*

- 1: /cgn2_6/ptodata/2/paa/PCTUS_COMB.pcp.*
- 2: /cgn2_6/ptodata/2/paa/US06_COMB.pcp.*
- 3: /cgn2_6/ptodata/2/paa/US07_COMB.pcp.*
- 4: /cgn2_6/ptodata/2/paa/US08_COMB.pcp.*
- 5: /cgn2_6/ptodata/2/paa/US081_COMB.pcp.*
- 6: /cgn2_6/ptodata/2/paa/US082_COMB.pcp.*
- 7: /cgn2_6/ptodata/2/paa/US083_COMB.pcp.*
- 8: /cgn2_6/ptodata/2/paa/US084_COMB.pcp.*
- 9: /cgn2_6/ptodata/2/paa/US085_COMB.pcp.*
- 10: /cgn2_6/ptodata/2/paa/US086_COMB.pcp.*
- 11: /cgn2_6/ptodata/2/paa/US087_COMB.pcp.*
- 12: /cgn2_6/ptodata/2/paa/US088_COMB.pcp.*
- 13: /cgn2_6/ptodata/2/paa/US089_COMB.pcp.*
- 14: /cgn2_6/ptodata/2/paa/US090_COMB.pcp.*
- 15: /cgn2_6/ptodata/2/paa/US091_COMB.pcp.*
- 16: /cgn2_6/ptodata/2/paa/US092_COMB.pcp.*
- 17: /cgn2_6/ptodata/2/paa/US093_COMB.pcp.*
- 18: /cgn2_6/ptodata/2/paa/US094_COMB.pcp.*
- 19: /cgn2_6/ptodata/2/paa/US095_COMB.pcp.*
- 20: /cgn2_6/ptodata/2/paa/US096_COMB.pcp.*
- 21: /cgn2_6/ptodata/2/paa/US097_COMB.pcp.*
- 22: /cgn2_6/ptodata/2/paa/US098_COMB.pcp.*
- 23: /cgn2_6/ptodata/2/paa/US099_COMB.pcp.*
- 24: /cgn2_6/ptodata/2/paa/US100_COMB.pcp.*
- 25: /cgn2_6/ptodata/2/paa/US101_COMB.pcp.*
- 26: /cgn2_6/ptodata/2/paa/US102_COMB.pcp.*
- 27: /cgn2_6/ptodata/2/paa/US60_COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8073	100.0	1596	1 PCT-US97-06830-4	Sequence 4, Appli
2	8073	100.0	1596	10 US-08-665-401-4	Sequence 4, Appli
3	8073	100.0	1596	13 US-08-978-277A-4	Sequence 4, Appli
4	8073	100.0	1596	23 US-09-502-432-4	Sequence 4, Appli
5	5421.5	67.2	1346	1 PCT-US97-06830-2	Sequence 2, Appli
6	5421.5	67.2	1346	10 US-08-665-401-2	Sequence 2, Appli

7	5421.5	67.2	1346	13	US-08-978-277A-2	Sequence 2, Appli
8	5421.5	67.2	1346	23	US-09-902-432-2	Sequence 2, Appli
9	4318.5	53.5	1782	27	US-60-389-987-2097	Sequence 2097, Ap
10	4318.5	53.5	1782	27	US-60-412-418-2097	Sequence 2097, Ap
11	4294	53.2	1781	21	US-09-738-877-3	Sequence 3, Appli
12	4294	53.2	1781	23	US-09-961-403-13	Sequence 13, Appli
13	4262.5	52.8	1795	1	PCT-US01-08631-51377	Sequence 51377, A
14	4253	52.7	1779	27	US-60-245-228-307	Sequence 307, App
15	780	9.7	732	21	US-09-758-442-562	Sequence 562, App
16	780	9.7	732	26	US-10-217-623-562	Sequence 562, App
17	752	9.3	212	27	US-60-182-569-1312	Sequence 1312, Ap
18	642	8.0	2768	20	US-09-614-150-31983	Sequence 31983, A
19	642	8.0	2768	27	US-60-173-464-24710	Sequence 24710, A
20	642	8.0	2768	27	US-60-191-637-31555	Sequence 31555, A
21	610.5	7.6	174	27	US-60-182-093-2249	Sequence 2249, Ap
22	561	6.9	194	27	US-60-186-662-1015	Sequence 1015, Ap
23	557.5	6.9	6358	27	US-60-191-637-27445	Sequence 27445, A
24	557.5	6.9	6358	27	US-60-191-681-22054	Sequence 22054, A
25	557.5	6.9	6815	20	US-09-614-150-27225	Sequence 27225, A
26	541.5	6.7	1770	27	US-60-196-710-5536	Sequence 5536, Ap
27	536	6.6	144	27	US-60-178-307-2458	Sequence 2458, Ap
28	534.5	6.6	169	27	US-60-196-710-6834	Sequence 6834, Ap
29	493	6.1	3924	19	US-09-538-092-1246	Sequence 1246, Ap
30	493	6.1	2468	19	US-09-538-092-1135	Sequence 1135, Ap
31	492	6.1	2468	19	US-09-538-092-1135	Sequence 1135, Ap
32	492	6.1	2468	23	US-09-976-594-726	Sequence 726, App
33	492	6.1	2468	27	US-60-240-409-726	Sequence 726, App
34	489	6.1	2519	1	PCT-US01-08631-46995	Sequence 46995, A
35	489	6.1	6642	27	US-60-360-039-5013	Sequence 5013, Ap
36	485	6.0	3263	20	US-09-614-150-28422	Sequence 28422, A
37	485	6.0	3267	27	US-60-191-637-27442	Sequence 27442, A
38	485	6.0	3267	27	US-60-191-681-22051	Sequence 22051, A
39	484	6.0	178	27	US-60-188-162-4950	Sequence 4950, Ap
40	473.5	5.9	8991	11	US-08-714-741-32	Sequence 32, Appli
41	470	5.8	2364	26	US-10-205-331-66	Sequence 66, Appli
42	465.5	5.8	212	27	US-60-196-710-4143	Sequence 4143, Ap
43	464.5	5.8	2024	27	US-60-167-217-5735	Sequence 5735, Ap
44	464.5	5.8	2024	27	US-60-173-464-4625	Sequence 4625, Ap
45	462	5.7	1786	13	US-08-973-462-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
PCT-US97-06830-4
; Sequence 4, Application PC/TUS9706830
; GENERAL INFORMATION:
; APPLICANT: Gelman, Irwin H.
; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10112-0228
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/06830
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/665,401
; FILING DATE: 18-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Richard S
; REGISTRATION NUMBER: 26,154

```

REFERENCE/DOCKET NUMBER: A30558 - 165/34008
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-408-2558
TELEFAX: 212-765-2519
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1596 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHEICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
PCT-US97-06830-4

Query Match      100.0%; Score 8073; DB 1; Length 1596;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1596; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGAGSTORSEPOEPAGSDTPSELVLSGHGPAEASGAAGDPADADPATKLPQKNGQLSS 60
DB 1 MGAGSTORSEPOEPAGSDTPSELVLSGHGPAEASGAAGDPADADPATKLPQKNGQLSS 60
QY 61 VNGVAEGDVAHVEENOGQEEVVDVQRESEEDVREKDRVEEMANSTAVEDITKDG 120
DB 61 VNGVAEGDVAHVEENOGQEEVVDVQRESEEDVREKDRVEEMANSTAVEDITKDG 120
QY 121 QETSEIIEQIPASBNNEEMVQPAESQANDVGFKKVFQFVKTKVDKNEKSDTVQL 180
DB 121 QETSEIIEQIPASBNNEEMVQPAESQANDVGFKKVFQFVKTKVDKNEKSDTVQL 180
QY 181 LTVKDEGGAASVAGAGDHQPSVETAVGESASKESEKOSTERKQEGTLKQEGSTEIP 240
DB 181 LTVKDEGGAASVAGAGDHQPSVETAVGESASKESEKOSTERKQEGTLKQEGSTEIP 240
QY 241 LQAESDOAEERAKDEGEKQKEPTKSPSPSPVNSSETTSSFKKFTTHGAMRKKTSS 300
DB 241 LQAESDOAEERAKDEGEKQKEPTKSPSPSPVNSSETTSSFKKFTTHGAMRKKTSS 300
QY 301 FKSKEDDLETKKKEQAEKVDDEBEKKTTPASEEGEPADDTQARLSAIVEKVELPL 360
DB 301 FKSKEDDLETKKKEQAEKVDDEBEKKTTPASEEGEPADDTQARLSAIVEKVELPL 360
QY 361 EDQVDDLEASSEKCAPLATEVFDEKMEAHQEVVAEVAHSTVEKTEEBEGGGGGAEGV 420
DB 361 EDQVDDLEASSEKCAPLATEVFDEKMEAHQEVVAEVAHSTVEKTEEBEGGGGGAEGV 420
QY 421 VEGTGESLPEPEKLAEPQVQAEPAEELMKSRMVCVSGDHTQULTDLSPEEKTLPKHPE 480
DB 421 VEGTGESLPEPEKLAEPQVQAEPAEELMKSRMVCVSGDHTQULTDLSPEEKTLPKHPE 480
QY 481 GIVSEVEMLSQERIKVQGSPLKLLFSSSGLKLLSGKKQKGRGGGGDEPEYOHINTE 540
DB 481 GIVSEVEMLSQERIKVQGSPLKLLFSSSGLKLLSGKKQKGRGGGGDEPEYOHINTE 540
QY 541 SPESADEOKGESSASPEPEETTCLEKPLAPODGAEEGTSDGEEKREGITPMASF 600
DB 541 SPESADEOKGESSASPEPEETTCLEKPLAPODGAEEGTSDGEEKREGITPMASF 600
QY 601 KKMVTPKKVRPSPSSDKEELEKVKASATLSTSTVSEMODEVKTVEGEOKPEEBKRV 660
DB 601 KKMVTPKKVRPSPSSDKEELEKVKASATLSTSTVSEMODEVKTVEGEOKPEEBKRV 660
QY 661 DTSVSEBALICVSSKKARAKASSSDDGCGPRTLGGDSHRAEASDKDAGDAVASTQ 720
DB 661 DTSVSEBALICVSSKKARAKASSSDDGCGPRTLGGDSHRAEASDKDAGDAVASTQ 720
QY 721 EDDOAGSSSPSPAGSPSGEGVSTWESFKRLVTPKKSLSLEKAEADSVSEQSTETE 780
DB 721 EDDOAGSSSPSPAGSPSGEGVSTWESFKRLVTPKKSLSLEKAEADSVSEQSTETE 780

```

```

QY 781 PSRESWSIKKFIPGRKKRADKQOATVDSGPVEINEDDPNVAVPLSFYNAVER 840
DB 781 PSRESWSIKKFIPGRKKRADKQOATVDSGPVEINEDDPNVAVPLSFYNAVER 840
QY 841 EKMAQGNTELPOLLGAVVYSEELSKTLVHTVSAVAVDGTGRAVSVSESPMSIASVTE 900
DB 841 EKMAQGNTELPOLLGAVVYSEELSKTLVHTVSAVAVDGTGRAVSVSESPMSIASVTE 900
QY 901 PLEHTAGAMPVEEVEKDIIAETPVLQTLPEGDAHDMDVTSVSDPTSEAVTATET 960
DB 901 PLEHTAGAMPVEEVEKDIIAETPVLQTLPEGDAHDMDVTSVSDPTSEAVTATET 960
QY 961 SEARTEVEVTSAGAEETTDVNSAVSQLTDSPTTEATPVQEVESGLTTEEBERQQA 1020
DB 961 SEARTEVEVTSAGAEETTDVNSAVSQLTDSPTTEATPVQEVESGLTTEEBERQQA 1020
QY 1021 ILQAVADKVEESQVPAQTQVQRTGSKALEKVEEEDSEVLASEKEDVMPKGPVQAG 1080
DB 1021 ILQAVADKVEESQVPAQTQVQRTGSKALEKVEEEDSEVLASEKEDVMPKGPVQAG 1080
QY 1081 AEHLAQSETOATPESLVEPEVTADVAVATCQVIKIQIMEQAVAPESSETLTDSETN 1140
DB 1081 AEHLAQSETOATPESLVEPEVTADVAVATCQVIKIQIMEQAVAPESSETLTDSETN 1140
QY 1141 GSTPLADSDTADGTQOQDETIDSQSKATAAVROSQVTEBEAATQKPEPSTLPNNVPAQE 1200
DB 1141 GSTPLADSDTADGTQOQDETIDSQSKATAAVROSQVTEBEAATQKPEPSTLPNNVPAQE 1200
QY 1201 EHGEPEGRDVLPTQOELTAAVAVLAKTEVQGEVQEDMDLGEKVEEVEFVHSGPNSQ 1260
DB 1201 EHGEPEGRDVLPTQOELTAAVAVLAKTEVQGEVQEDMDLGEKVEEVEFVHSGPNSQ 1260
QY 1261 KAADVTYDSEVMYAGCQEKESTEVQSLSLEGEMETDVEKREKTRPEQVSEGEQETA 1320
DB 1261 KAADVTYDSEVMYAGCQEKESTEVQSLSLEGEMETDVEKREKTRPEQVSEGEQETA 1320
QY 1321 APHEBGTGKRVLTLDMPSSRGKALGSLGSLPPODRAKGCIEVOYQSLDTVTYQTAE 1380
DB 1321 APHEBGTGKRVLTLDMPSSRGKALGSLGSLPPODRAKGCIEVOYQSLDTVTYQTAE 1380
QY 1381 AVEKVIETVVISSETGESPECGAHLPLPAKSSATGCHMTLOHADTVPLGPGSOAESIP 1440
DB 1381 AVEKVIETVVISSETGESPECGAHLPLPAKSSATGCHMTLOHADTVPLGPGSOAESIP 1440
QY 1441 IVPAPESLHPDLQGEISASQRESEBEDKPDAGPDAGKESSTAIEKVLKAEDEILLE 1500
DB 1441 IVPAPESLHPDLQGEISASQRESEBEDKPDAGPDAGKESSTAIEKVLKAEDEILLE 1500
QY 1501 SKSNKIYVNVQTVANDQPARFETAPETHAYDSOTQOVPAKRLDSREPRCWTMKDARKKH 1560
DB 1501 SKSNKIYVNVQTVANDQPARFETAPETHAYDSOTQOVPAKRLDSREPRCWTMKDARKKH 1560
QY 1561 PVPOREDLOVLTVEAMAOPRKCLPRLQKAPVSK 1596
DB 1561 PVPOREDLOVLTVEAMAOPRKCLPRLQKAPVSK 1596

```

```

RESULT 2
US-08-665-401-4
Sequence 4, Application US/08665401
GENERAL INFORMATION:
APPLICANT: Gelman, Irwin H.
TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESS: Brumbaugh, Graves, Donohue & Raymond
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10112-0228
COMPUTER READABLE FORM:

```

```

; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/665,401
; FILING DATE: 18-JUN-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Richard S
; REGISTRATION NUMBER: 26,154
; REFERENCE/DOCKET NUMBER: A30558 - 165/34008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-408-2558
; TELEFAX: 212-765-2519
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1596 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; US-08-665-401-4

```

```

Query Match 100.0%; Score 8073; DB 10; Length 1596;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1596; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGAGSSTEQSPPEAGSDTSELVLSHGPAEASGAAGDPADPATKLPQKNGQLSS 60
Db 1 MGAGSSTEQSPPEAGSDTSELVLSHGPAEASGAAGDPADPATKLPQKNGQLSS 60

Qy 61 VNGVAEQGDVHVQENQBGQEEVVDVGQRESEDVREKDRVEMAANSTAVEDITKDG 120
Db 61 VNGVAEQGDVHVQENQBGQEEVVDVGQRESEDVREKDRVEMAANSTAVEDITKDG 120

Qy 121 QEETSEIIQIPASENNVEMVQPAESQANDVGFKVFKFVGFKFTVKDKNEKSDTVQL 180
Db 121 QEETSEIIQIPASENNVEMVQPAESQANDVGFKVFKFVGFKFTVKDKNEKSDTVQL 180

Qy 181 LTVKDEGEAGASVAGADHQPSPVETAVGESASKESELKOSTEKQEGTLKQEGSSTEIP 240
Db 181 LTVKDEGEAGASVAGADHQPSPVETAVGESASKESELKOSTEKQEGTLKQEGSSTEIP 240

Qy 241 LQASDQAAEAREKDEGEKEKPTKSPSPSPVNSSETTSSPKFFTHGWAGHRKKT 300
Db 241 LQASDQAAEAREKDEGEKEKPTKSPSPSPVNSSETTSSPKFFTHGWAGHRKKT 300

Qy 301 FKSKEDDLETAEKREKQEAQKVEEKEKTEPASEEQAEDTDQARLSADYKVELPL 360
Db 301 FKSKEDDLETAEKREKQEAQKVEEKEKTEPASEEQAEDTDQARLSADYKVELPL 360

Qy 361 EDQVGDLEASSEEKCAPLATEVFDEKMAHQEVVAEVHVSHTVEKTEEPQGGGGAEGGW 420
Db 361 EDQVGDLEASSEEKCAPLATEVFDEKMAHQEVVAEVHVSHTVEKTEEPQGGGGAEGGW 420

Qy 421 VEGTGESLPPEKLAEPQVPOEAPAEELMKSRMCVSGGDHTQLTDLSPREKTLPKHPE 480
Db 421 VEGTGESLPPEKLAEPQVPOEAPAEELMKSRMCVSGGDHTQLTDLSPREKTLPKHPE 480

Qy 481 GIVSEVEMLSQERIKVQGSPLKLLFSSSGGLKLSGKKQKRGGGGDEEPEGYQHIHTE 540
Db 481 GIVSEVEMLSQERIKVQGSPLKLLFSSSGGLKLSGKKQKRGGGGDEEPEGYQHIHTE 540

Qy 541 SPESADEQKGESSASSPEPETTCLKGPLEAPQDGAEBEGTTSDGKKREGITPWASF 600

```

RESULT 3

```

Db 541 SPESADEQKGESSASSPEPETTCLKGPLEAPQDGAEBEGTTSDGKKREGITPWASF 600
Qy 601 KKMVTPKKRVRRPSESDEKELEKVKSATLSSSTDSTVSEMODEVKTGVEEQKPEEPKRV 660
Db 601 KKMVTPKKRVRRPSESDEKELEKVKSATLSSSTDSTVSEMODEVKTGVEEQKPEEPKRV 660
Qy 661 DTSVSWEALI:CVGSSKKRARKASSDDGGPRTLGGDSHRAEASKOKKAGTDAVPASTQ 720
Db 661 DTSVSWEALI:CVGSSKKRARKASSDDGGPRTLGGDSHRAEASKOKKAGTDAVPASTQ 720
Qy 721 EDOAQGSSSPPEAGSPSEGEVSTWESFKRLVTPPKSKSKLEKAEADSVSEOLSTEIE 780
Db 721 EDOAQGSSSPPEAGSPSEGEVSTWESFKRLVTPPKSKSKLEKAEADSVSEOLSTEIE 780
Qy 781 PSREESWVSIKKFIPGRRKKRADGKQQAATVEDSGPVEINEDDDPNVPAVPLSYNAVER 840
Db 781 PSREESWVSIKKFIPGRRKKRADGKQQAATVEDSGPVEINEDDDPNVPAVPLSYNAVER 840
Qy 841 EKMEAQGNTELPLLGAIVYVSEELSKTLVHTVSVAVIDGTRAVTSVEERSPSWISASVTE 900
Db 841 EKMEAQGNTELPLLGAIVYVSEELSKTLVHTVSVAVIDGTRAVTSVEERSPSWISASVTE 900
Qy 901 PLEHTAGEAMPVEEVTEKDI:IAEETPVLTQTLPKGDAHDDMVTSEVDFTSSEAVTATET 960
Db 901 PLEHTAGEAMPVEEVTEKDI:IAEETPVLTQTLPKGDAHDDMVTSEVDFTSSEAVTATET 960
Qy 961 SEALRTEEVTEASGAETTDWVSAVSQLTDSPTTEATPVQEVESGVLDTTEEROTQA 1020
Db 961 SEALRTEEVTEASGAETTDWVSAVSQLTDSPTTEATPVQEVESGVLDTTEEROTQA 1020
Qy 1021 ILQAVADKVKESQVPATQTVQRTGSKALEKVEEVEDSEVLASEKEKDMVMPKGPVQBSAG 1080
Db 1021 ILQAVADKVKESQVPATQTVQRTGSKALEKVEEVEDSEVLASEKEKDMVMPKGPVQBSAG 1080
Qy 1081 AEHLAQSEGTQATPESLEVPETADVHDVATCOVI:KLQQLMEQAVAPESSETLTDSETN 1140
Db 1081 AEHLAQSEGTQATPESLEVPETADVHDVATCOVI:KLQQLMEQAVAPESSETLTDSETN 1140
Qy 1141 GSTPLADSDTAGTQDQDETIDSQDSKATAAVRSQVTEEAATAQKGEPSLTPNNVPAQE 1200
Db 1141 GSTPLADSDTAGTQDQDETIDSQDSKATAAVRSQVTEEAATAQKGEPSLTPNNVPAQE 1200
Qy 1201 EHGEPEGRDVLPTQQLTAAAVPVLAKTEVQGEVGDWLDGKVKEEQEVFVHSGPNSQ 1260
Db 1201 EHGEPEGRDVLPTQQLTAAAVPVLAKTEVQGEVGDWLDGKVKEEQEVFVHSGPNSQ 1260
Qy 1261 KAADVTYDSEVMGVAGQCKEKESTEVSQSLSEEGEMETDVEKEKETPKPEQVSEEGQETA 1320
Db 1261 KAADVTYDSEVMGVAGQCKEKESTEVSQSLSEEGEMETDVEKEKETPKPEQVSEEGQETA 1320
Qy 1321 APHEGTYGKPVLTLDMPSSERKALGSLGSSPSLPDQDAGCIEVOVQSLDTTDTTQTAE 1380
Db 1321 APHEGTYGKPVLTLDMPSSERKALGSLGSSPSLPDQDAGCIEVOVQSLDTTDTTQTAE 1380
Qy 1381 AVEKVIETVWISGETSPCEVGAHLLPAEKSSATGCGHWTLQHAEDTVPPLGPESQAEIPI 1440
Db 1381 AVEKVIETVWISGETSPCEVGAHLLPAEKSSATGCGHWTLQHAEDTVPPLGPESQAEIPI 1440
Qy 1441 IVTPAPESTLHPDLQGEISASQORRSSEEDKPDAGPDADGKESTAI:EKVLKAEPEIILELE 1500
Db 1441 IVTPAPESTLHPDLQGEISASQORRSSEEDKPDAGPDADGKESTAI:EKVLKAEPEIILELE 1500
Qy 1501 SKSNKIVLNIQTAVDQFARTETAPETHAVDSQTPACRLDSREPNCRTMKDKAKMKH 1560
Db 1501 SKSNKIVLNIQTAVDQFARTETAPETHAVDSQTPACRLDSREPNCRTMKDKAKMKH 1560
Qy 1561 PVQOPREDLQVLTVLEAWAQPKCLPRLQLKAPVSK 1596
Db 1561 PVQOPREDLQVLTVLEAWAQPKCLPRLQLKAPVSK 1596

```

US-08-978-277A-4
; Sequence 4, Application US/08978277A
; GENERAL INFORMATION:
; APPLICANT: Gelman, Irwin H.
; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10112-0228
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: IBM Compatible
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/978,277A
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/665,401
; FILING DATE: 18-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Richard S
; REGISTRATION NUMBER: 26,154
; REFERENCE/DOCKET NUMBER: A30558 - 165/34008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-408-2558
; TELEFAX: 212-765-2519
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1596 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; US-08-978-277A-4
Query Match 100.0%; Score 8073; DB 13; Length 1596;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1596; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 361 EDVQDLEASSEKCAPLATVEPDEKMEAHQEVYAEVHVSTVEKTEBEEQGGGABEGGV 420
Db 361 EDVQDLEASSEKCAPLATVEPDEKMEAHQEVYAEVHVSTVEKTEBEEQGGGABEGGV 420
QY 421 VEGTGESLPPEKLAPOEVPQEAPEAEIMKSRMCVSGGDHQTJLSPREKTLPKPHE 480
Db 421 VEGTGESLPPEKLAPOEVPQEAPEAEIMKSRMCVSGGDHQTJLSPREKTLPKPHE 480
QY 481 GIVSEVEMLSQERIKYQGSPLKTLFSSSGIKKLSGKKQKRGGGGDEEPGEYOHITHTE 540
Db 481 GIVSEVEMLSQERIKYQGSPLKTLFSSSGIKKLSGKKQKRGGGGDEEPGEYOHITHTE 540
QY 541 SPESADEQKSSSASPEPEETTCLEKGPLAEODEAEAGCTTSDCKREGITTPMASF 600
Db 541 SPESADEQKSSSASPEPEETTCLEKGPLAEODEAEAGCTTSDCKREGITTPMASF 600
QY 601 KKMVTPKKRVRRPSESDKEELEKYKSATLSTSTVSEMODEVKTVEEQKPREPKRV 660
Db 601 KKMVTPKKRVRRPSESDKEELEKYKSATLSTSTVSEMODEVKTVEEQKPREPKRV 660
QY 661 DTSVSWELICVSSKKRARKASSDDEGCPRTLGDSSHRAEASKDEACTDAVPASTQ 720
Db 661 DTSVSWELICVSSKKRARKASSDDEGCPRTLGDSSHRAEASKDEACTDAVPASTQ 720
QY 721 EODQAQSSSPREPAGSPSEGGCVSTWESFKRLVTPRKKSKSLEKAKDSVBOULSTIE 780
Db 721 EODQAQSSSPREPAGSPSEGGCVSTWESFKRLVTPRKKSKSLEKAKDSVBOULSTIE 780
QY 781 PSRESVWSIKKFIPIGRKKRADKQEOATVEDSGPVEINDDPNVAVVPLSEYNAYER 840
Db 781 PSRESVWSIKKFIPIGRKKRADKQEOATVEDSGPVEINDDPNVAVVPLSEYNAYER 840
QY 841 EKMEAQNTLPLQLGAVYSEELSKTLVHTVSAVIDGTRAVTSVEERSPSWISASYTE 900
Db 841 EKMEAQNTLPLQLGAVYSEELSKTLVHTVSAVIDGTRAVTSVEERSPSWISASYTE 900
QY 901 PLEHTAGAMPPEVEVETKDIIAETEPVLTOTLPEGDAHDMDTSEDPFSEAVTATET 960
Db 901 PLEHTAGAMPPEVEVETKDIIAETEPVLTOTLPEGDAHDMDTSEDPFSEAVTATET 960
QY 961 SEALTREEVETASGAETTDWASVSOULTSPDTEATPVQEVESGVLDTSEERQTOA 1020
Db 961 SEALTREEVETASGAETTDWASVSOULTSPDTEATPVQEVESGVLDTSEERQTOA 1020
QY 1021 ILQAVADKVKESQVPATQVQTGTSKALEKVEEVEDSEVLASEKEDVMPKGPVQEG 1080
Db 1021 ILQAVADKVKESQVPATQVQTGTSKALEKVEEVEDSEVLASEKEDVMPKGPVQEG 1080
QY 1081 AEHLAQSSETCOATPESLEVEVEVADVADVATCOVYIKLQQLMEQAVAPESSETLTDSETN 1140
Db 1081 AEHLAQSSETCOATPESLEVEVEVADVADVATCOVYIKLQQLMEQAVAPESSETLTDSETN 1140
QY 1141 GSTPLABSDTADGTQOQDETTIDSQSKATAVROSOVTEEBAAATQOKEPSTLPNNVPOE 1200
Db 1141 GSTPLABSDTADGTQOQDETTIDSQSKATAVROSOVTEEBAAATQOKEPSTLPNNVPOE 1200
QY 1201 EHGEPEGRDVLPTQOELTAAAVPVLAKTEBVGQGEVDMLDGEKVKESQEVFVHSGPNSQ 1260
Db 1201 EHGEPEGRDVLPTQOELTAAAVPVLAKTEBVGQGEVDMLDGEKVKESQEVFVHSGPNSQ 1260
QY 1261 KAADVTYDSEVMYAGCOEKESTEVQSLGLEEGMETDVEKEKETPPEQVSEBGEQETA 1320
Db 1261 KAADVTYDSEVMYAGCOEKESTEVQSLGLEEGMETDVEKEKETPPEQVSEBGEQETA 1320
QY 1321 APEHGTYGKPVLTLDMPSSRGKALSGSPSLPPODKXGCJEVOYQSLDITVTYTOAE 1380
Db 1321 APEHGTYGKPVLTLDMPSSRGKALSGSPSLPPODKXGCJEVOYQSLDITVTYTOAE 1380
QY 1381 AVEKVIETVVISETGESPEECVGAHLIPAEKSSATGHWTLQHAEDTVPLGPESQAESIPI 1440
Db 1381 AVEKVIETVVISETGESPEECVGAHLIPAEKSSATGHWTLQHAEDTVPLGPESQAESIPI 1440

Qy 1441 IVTPAPESTLHPDLQGEISASQSRSEEDKPDAGPDADGKESTAIKVKLKAEPILILE 1500
Db 1441 IVTPAPESTLHPDLQGEISASQSRSEEDKPDAGPDADGKESTAIKVKLKAEPILILE 1500
Qy 1501 SKSNKIVLNVITQAVDQFARTETAPETHAYDSQTVQVPACRLDSREPNRCWTMCKDAKMKH 1560
Db 1501 SKSNKIVLNVITQAVDQFARTETAPETHAYDSQTVQVPACRLDSREPNRCWTMCKDAKMKH 1560
Qy 1561 VVPQPREDLQVLTVLEAWAQPRLCLPRQLKAPVSK 1596
Db 1561 VVPQPREDLQVLTVLEAWAQPRLCLPRQLKAPVSK 1596

RESULT 4
US-09-902-432-4
; Sequence 4, Application US/09902432
; GENERAL INFORMATION:
; APPLICANT: Irwin H. Gelman
; APPLICANT: Susan G. Jaken
; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
; FILE REFERENCE: A30558-A-FWC-A, 070156.0597
; CURRENT APPLICATION NUMBER: US/09/902.432
; CURRENT FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 08/978,277
; PRIOR FILING DATE: 1997-11-25
; PRIOR APPLICATION NUMBER: 08/665,401
; PRIOR FILING DATE: 1996-06-18
; PRIOR APPLICATION NUMBER: 08/635,121
; PRIOR FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1596
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-902-432-4

Query Match 100.0%; Score 8073; DB 23; Length 1596;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1596; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGASSTEQRSPEQAGSDTPELVLVSHGPAEASGAGNAGPADADPATKLQKNGQLSS 60
Db 1 MGASSTEQRSPEQAGSDTPELVLVSHGPAEASGAGNAGPADADPATKLQKNGQLSS 60
Qy 61 VNGVAEQGDVHVQENQGEQEEVVDVQRESEDVREKDRVEEMANSTAVEDITKDG 120
Db 61 VNGVAEQGDVHVQENQGEQEEVVDVQRESEDVREKDRVEEMANSTAVEDITKDG 120
Qy 121 QEETSEIIIEQIPASNNVVEWQPAESQANDVGFKKVPKFKFTVKCKDKNEKSDTVQL 180
Db 121 QEETSEIIIEQIPASNNVVEWQPAESQANDVGFKKVPKFKFTVKCKDKNEKSDTVQL 180
Qy 181 LTVKKDEGEAGASVAGDHDQPSVETAVGSAKSESELKQSTKEQGTLLKQBOSSSTEIP 240
Db 181 LTVKKDEGEAGASVAGDHDQPSVETAVGSAKSESELKQSTKEQGTLLKQBOSSSTEIP 240
Qy 241 LQAESDQAAEAEKDEGEKEKEPTKSPESPSPVNSSETTSSPKPPTHGAWGRKKT 300
Db 241 LQAESDQAAEAEKDEGEKEKEPTKSPESPSPVNSSETTSSPKPPTHGAWGRKKT 300
Qy 301 FKSKEDDLETAEKRKQAEKVDDEEKEKTEPASEEQEPAEDTDQARLSADYKVELPL 360
Db 301 FKSKEDDLETAEKRKQAEKVDDEEKEKTEPASEEQEPAEDTDQARLSADYKVELPL 360
Qy 361 EDQVGDLEASSEKCAPLATVFEKMEAHQEVVAEVHVSUVTEKTEBERQGGGGAEGGV 420
Db 361 EDQVGDLEASSEKCAPLATVFEKMEAHQEVVAEVHVSUVTEKTEBERQGGGGAEGGV 420
Qy 421 VEGTGESLPPKELAPQVPOEAPAEELMKSRMCVSGGDHTQLTDLSPBEKTLPKHPE 480
Db 421 VEGTGESLPPKELAPQVPOEAPAEELMKSRMCVSGGDHTQLTDLSPBEKTLPKHPE 480

Qy 481 GIYSEVEMLSQBERIKVQGSPLKKLFSSSGLKKLSGKKQKGRGGGDEBPGEYQHITHTE 540
Db 481 GIYSEVEMLSQBERIKVQGSPLKKLFSSSGLKKLSGKKQKGRGGGDEBPGEYQHITHTE 540
Qy 541 SPESADQKGESSASSPEEPETTCLEKGPLEAPQDGEAEEGTTSDGCKKREGITPWASF 600
Db 541 SPESADQKGESSASSPEEPETTCLEKGPLEAPQDGEAEEGTTSDGCKKREGITPWASF 600
Qy 601 KKMVTPKRRVRRPESDKSEBELEKVKSATLSSDSTVSEMQDEVKTVGEEKQKPEEPKRRV 660
Db 601 KKMVTPKRRVRRPESDKSEBELEKVKSATLSSDSTVSEMQDEVKTVGEEKQKPEEPKRRV 660
Qy 661 DTSVWEALICVSSSKRKARKASSDDDEGPRTLGGDSHRAEASKOKERAGTDVAPASTQ 720
Db 661 DTSVWEALICVSSSKRKARKASSDDDEGPRTLGGDSHRAEASKOKERAGTDVAPASTQ 720
Qy 721 EQDAQGSSSPPEAGSPSEGEVSTWESFKRLVTPRKKSLSKLEEKAEADSVEQLSTEIE 780
Db 721 EQDAQGSSSPPEAGSPSEGEVSTWESFKRLVTPRKKSLSKLEEKAEADSVEQLSTEIE 780
Qy 781 PSREESWVSICKTIPGRRKKRKGQKQATVEDSGPVEINEDDPNVPAVPLSEYNAVER 840
Db 781 PSREESWVSICKTIPGRRKKRKGQKQATVEDSGPVEINEDDPNVPAVPLSEYNAVER 840
Qy 841 EKMEAQGNTELPQLLGAVVYSEBELSKTLVHTVSVAVTDGTRAVTSVEERSPSWISASVTE 900
Db 841 EKMEAQGNTELPQLLGAVVYSEBELSKTLVHTVSVAVTDGTRAVTSVEERSPSWISASVTE 900
Qy 901 PLEHTAGEAMPVVEEVTEKDI IAEETPVLTQTLPEGKAHDDMVTSVDFTSEAVTATET 960
Db 901 PLEHTAGEAMPVVEEVTEKDI IAEETPVLTQTLPEGKAHDDMVTSVDFTSEAVTATET 960
Qy 961 SEALRTEEVTEASGAETTTDMVSAVSQLTDSPTTTEATPVQEVESGVLDTTEEROTQA 1020
Db 961 SEALRTEEVTEASGAETTTDMVSAVSQLTDSPTTTEATPVQEVESGVLDTTEEROTQA 1020
Qy 1021 ILQAVADKVKESQVATQTVQRTGSKALEKVEEVEDSEVLASEKBDKVPKGPVQEGAG 1080
Db 1021 ILQAVADKVKESQVATQTVQRTGSKALEKVEEVEDSEVLASEKBDKVPKGPVQEGAG 1080
Qy 1081 AEHLAQGSETGOATPESLEVPETADVHVATCOVILKLOLMEOAVALPESSETLTDSETN 1140
Db 1081 AEHLAQGSETGOATPESLEVPETADVHVATCOVILKLOLMEOAVALPESSETLTDSETN 1140
Qy 1141 GSTPLADSDTAGTQDQDETIDSQDSKATAAVROSQVTEEEAATAQKEEPSTLPNNVPAQE 1200
Db 1141 GSTPLADSDTAGTQDQDETIDSQDSKATAAVROSQVTEEEAATAQKEEPSTLPNNVPAQE 1200
Qy 1201 EHGEPEGRDVLPTQOQELTAAAVPVVLAKTEVQGEVGDWLDGKVKBEQEVFVHSGPNSQ 1260
Db 1201 EHGEPEGRDVLPTQOQELTAAAVPVVLAKTEVQGEVGDWLDGKVKBEQEVFVHSGPNSQ 1260
Qy 1261 KAADVTYDSEVMGVAGCQEKESTEVQSLSEEGEMETDVEKEKETPEQVSEBGEQETA 1320
Db 1261 KAADVTYDSEVMGVAGCQEKESTEVQSLSEEGEMETDVEKEKETPEQVSEBGEQETA 1320
Qy 1321 APEHEGTGKPVLTLDMPSSERKALGSLGSPSLPDQDKAGCTIEVOVQSLDITVTTOTAE 1380
Db 1321 APEHEGTGKPVLTLDMPSSERKALGSLGSPSLPDQDKAGCTIEVOVQSLDITVTTOTAE 1380
Qy 1381 AVEKVIETVVISSETGESPECVGAHLLPAEKSSATGGHWTLOHAEDTVPLGPESQAESIP 1440
Db 1381 AVEKVIETVVISSETGESPECVGAHLLPAEKSSATGGHWTLOHAEDTVPLGPESQAESIP 1440
Qy 1441 IVTPAPESTLHPDLQGEISASQSRSEEDKPDAGPDADGKESTAIKVKLKAEPILILE 1500
Db 1441 IVTPAPESTLHPDLQGEISASQSRSEEDKPDAGPDADGKESTAIKVKLKAEPILILE 1500
Qy 1501 SKSNKIVLNVITQAVDQFARTETAPETHAYDSQTVQVPACRLDSREPNRCWTMCKDAKMKH 1560
Db 1501 SKSNKIVLNVITQAVDQFARTETAPETHAYDSQTVQVPACRLDSREPNRCWTMCKDAKMKH 1560
Qy 1561 VVPQPREDLQVLTVLEAWAQPRLCLPRQLKAPVSK 1596

Db 1561 PVQPREDLQVLTLEAWAQPRLCLPRLQLAPVSK 1596

RESULT 5

PCT-US97-06830-2

Sequence 2, Application PC/TUS9706830
 GENERAL INFORMATION:
 APPLICANT: Gelman, Irwin H.
 TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
 NUMBER OF SEQUENCES: 20
 CORRESPONDENCE ADDRESS:
 ADDRESS: Brumbaugh, Graves, Donohue & Raymond
 STREET: 30 Rockefeller Plaza
 CITY: New York
 STATE: NY
 COUNTRY: USA
 ZIP: 10112-0228
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq Version 1.5
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US97/06830
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/665,401
 FILING DATE: 18-JUN-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Clark, Richard S
 REGISTRATION NUMBER: 26,154
 REFERENCE/DOCKET NUMBER: A30558 - 165/34008
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-408-2558
 TELEFAX: 212-765-2519
 TELEX:
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1346 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 HYPOTHEetical: NO
 ANTI-SENSE: NO
 FRAGMENT TYPE: internal
 ORIGINAL SOURCE:
 PCT-US97-06830-2

Query Match 67.2%; Score 5421.5; DB 1; Length 1346;
 Best Local Similarity 91.6%; Pred. No. 4.9e-304; Indels 15; Gaps 7;
 Matches 1098; Conservative 24; Mismatches 62;

Qy 387 MEAHOEVAEVAHVSTVEKTEBEGGGGGEAAGGVVEGTESLPPEKLAPOEVPQEAEDA 446
 Db 1 MEAHOEVAEVAHVSTVEKTEBEGGGGGEAAGGVVEGTESLPPEKLAPOEVPQEAEDA 60
 Qy 447 EELMSKRENCVSGGHDITLDSPEKTLPKHPEGVSEVEMLSQERKLVGSSPLKCLF 506
 Db 61 EELMSKRENCVSGGHDITLDSPEKTLPKHPEGVSEVEMLSQERKLVGSSPLKCLF 120
 Qy 507 SSSGKIKLSGKKQKRGGGGDEPGEVQHITESPESADEQKGSASSPPEPTTCL 566
 Db 121 SSSGKIKLSGKKQKRGGGGDEPGEVQHITESPESADEQKGSASSPPEPTTCL 180
 Qy 567 EKGPLEAFODGAEFGTSDGKKRE---GITPMASFKAQVTPKKYVRRPESDKEBEL 622
 Db 181 EKGPLEAFPMGLRKRELRL-GEKKKDHSLGI-----LQKODTQETVARPESDKEBEL 234
 Qy 623 EKVKSATLSSDTSVSEMODEVYKVGEEQKPEEPKRVDTVSWEALIVGSSKKRARA 682

Db 235 EKVKSATLSSDTSVSEMODEVYKVGEEQKPEEPKRVDTVSWEALIVGSSKKRARA 294
 Qy 683 SSSDDEGGPRTLGGDSHRAEASKDEAGTDVAPASTQEDQAGSSSPPEPAGSPSEGG 742
 Db 295 SSSDIR-GPRTLGGGQSQNRQGRSRRTDAPVASTQEDQAGSSSPPEPAGSPSEGG 353
 Qy 743 VSTWESFRRLVTPKKSKSKLEEK-AEDSSVEQLSTIEPSRESNWS--IKKIPGRK 799
 Db 354 VSTWESFRRLVTPKKSKSKLEEKAGRTLVAGAGCPLRSNRKKNLGFPLRNSPPGGR 413
 Qy 800 KRADKQEOATVEDESGVEINEEDPNVAVVPLSEYNAVEREKKEAOGNTLPOLLAVY 859
 Db 414 KQNGROEQATVEDSGVEINEEDPDVAVVPLSEYNAVEREKKEAOGNTLPOLLAVY 472
 Qy 860 VSEELSKTLVHTVSVAVIDGTRAVTSVEERSPSWISASVTEPLEHTAGEAMPVEYTEK 919
 Db 473 VSEELSKTLVHTVSVAVIDGTRAVTSVEERSPSWISASVTEPLEHTAGEAMPVEYTEK 532
 Qy 920 DIIAETPVLTQTLPEKGDADDMVTSEVDPTSRAVATETSEALRTEEYTEASGARETT 979
 Db 533 DIIAETPVLTQTLPEKGDADDMVTSEVDPTSRAVATETSEALRTEEYTEASGARETT 592
 Qy 980 DMVAVSQLTDSPTTEBATPVQEVESGLDTEEEERQTOAILQAVADKVEESQVPATQ 1039
 Db 593 DMVAVSQLTDSPTTEBATPVQEVESGLDTEEEERQTOAILQAVADKVEESQVPATQ 652
 Qy 1040 TVQRTGSKALEKEVEBEDSEVLASEKEKDMVPMKPYQOEAQAEHLAOGSETQATPESLE 1099
 Db 653 TVQRTGSKALEKEVEBEDSEVLASEKEKDMVPMKPYQOEAQAEHLAOGSETQATPESLE 712
 Qy 1100 VPEVTADVDVATQVYIKLQQLMEQAVAPESSETLTTSETNGSTPLADSDTADTQDET 1159
 Db 713 VPEVTADVDVATQVYIKLQQLMEQAVAPESSETLTTSETNGSTPLADSDTADTQDET 772
 Qy 1160 IDSODSKATAVROSQVTEEBEATAQKEPSTLPNNVPAQEHGEERQVLEPTQOELT 1219
 Db 773 IDSODSKATAVROSQVTEEBEATAQKEPSTLPNNVPAQEHGEERQVLEPTQOELT 832
 Qy 1220 AAAPVVLAKTEVGEVDMWLDGKVKKEBOEVPVHSGPNSQKADVTYDSEVMVAGQOE 1279
 Db 833 AAAPVVMKTEVGEVDMWLDGKVKKEBOEVPVHSGPNSQKADVTYDSEVMVAGQOE 892
 Qy 1280 KESTEVQSLSEBEMETDVAKETKPEVYSEEGQETAPHEBCTYGRPVTLDMPS 1339
 Db 893 KESTEVQSLSEBEMETDVAKETKPEVYSEEGQETAPHEBCTYGRPVTLDMPS 952
 Qy 1340 SERGKALGSLGSPSLPDODVAGCIEVOVQSLDTTVYQTAFAVEKVIETVVISETGESPE 1399
 Db 953 SERGKALGSLGSPSLPDODVAGCIEVOVQSLDTTVYQTAFAVEKVIETVVISETGESPE 1012
 Qy 1400 CVGAHLPAEKSSATGGHMTLQHAEDTVPLGESQASIPITVPAPBSTLHPDLQGEIS 1459
 Db 1013 CVGAHLPAEKSSATGGHMTLQHAEDTVPLGESQASIPITVPAPBSTLHPDLQGEIS 1072
 Qy 1460 ASQRESEEBEKPDAGPADGKESTAIKYLKAPPELIELESKSNKYLVNIQTAVDOFA 1519
 Db 1073 ASQRESEEBEKPDAGPADGKESTAIKYLKAPPELIELESKSNKYLVNIQTAVDOFA 1132
 Qy 1520 RTEPAETHAVDSQTVQACGLDSREPNRCWTMMKDKMKHPVQOPREDLVLTVLEAW 1578
 Db 1133 RTEPAETHAVDSQTVQACGLDSREPNRCWTMMKDKMKHPVQOPREDLVLTVLEAW 1191

RESULT 6

US-08-665-401-2

Sequence 2, Application US/08665401
 GENERAL INFORMATION:
 APPLICANT: Gelman, Irwin H.
 TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
 NUMBER OF SEQUENCES: 20
 CORRESPONDENCE ADDRESS:
 ADDRESS: Brumbaugh, Graves, Donohue & Raymond
 STREET: 30 Rockefeller Plaza

; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10112-0228
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/665,401
; FILING DATE: 18-JUN-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Richard S
; REGISTRATION NUMBER: 26,154
; REFERENCE/DOCKET NUMBER: A30558 - 165/34008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-408-2558
; TELEFAX: 212-765-2519
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1346 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; US-08-665-401-2

Query Match 67.2%; Score 5421.5; DB 10; Length 1346;
Best Local Similarity 91.6%; Pred. No. 4.9e-304;
Matches 1098; Conservative 24; Mismatches 62; Indels 15; Gaps 7;
QY 387 MEAHQEVVAHVHSTVETKEEQQGGGAEAGGVVVEGTGSLPEKLAHQEVQAEPA 446
DB 1 MEAHQEVVAHVHSTVETKEEQQGGGAEAGGVVVEGTGSLPEKLAHQEVQAEPA 60
QY 447 BELKMSREMCVSGGDHTQLTDLSPPEKTLPKHPEGIVSEVEMLSQRIKVGQSPKLKLF 506
DB 61 BELKMSREMCVSGGDHTQLTDLSPPEKTLPKHPEGIVSEVEMLSQRIKVGQSPKLKLF 120
QY 507 SSSGLKLLSGKKQKRGGGGDEBPYQHHTSPESADEQKGESSASPEEPETTC 566
DB 121 SSSGLKLLSGKKQKRGGGGDEBPYQHHTSPESADEQKGESSASPEEPETTC 180
QY 567 EKGPLEAQDEAGEETTSDEKRE----GITWAFKMTVPKVRPRPSESDEKEEL 622
DB 181 EKGPLEAPRMGKLKELL-GEKKRKHSLGI-----LQKGDGTQETVRRPSESDEKEEL 234
QY 623 EKVSATLSSTDSVSEMODVKTVGEQKPEEPKRRVDTSVSWEALICVSSSKRKARKA 682
DB 235 EKVSATLSSTDSVSEMODVKTVGEQKPEEPKRRVDTSVSWEALICVSSSKRKARKA 294
QY 683 SSSDDEGGPRTLGGDSHRAEASKXDEAGTDVAPASTQEQDAQSSSPPEAGSPEG 742
DB 295 SSSDIR-GPRTLGGGQSQRGQQRSRDQVAPASTQEQDAQSSSPPEAGSPEG 353
QY 743 VSTWESFKRLVTPPKSKSKLEEK-ADDSVBEQLSTIEPSPRESWVS--IKKFIPORRK 799
DB 354 VSTWESFKRLVTPPKSKSKLEEK-ADDSVBEQLSTIEPSPRESWVS--IKKFIPORRK 413
QY 800 KRADKQEQATVSDGSPVEINEDDPNPVAVVPLSEYNVAREKMEAGNTELPQLLGAVY 859
DB 414 KGQMGREQATVSDGSPVEINEDDPNPVAVVPLSEYNVAREKMEAGNTELPQLLGAVY 472

QY 860 VSELSKTLVHTVSVAVIDGTRAVTSVEERSPSWISASVTEPLEHTAGEAMPPVEVTEK 919
DB 473 VSELSKTLVHTVSVAVIDGTRAVTSVEERSPSWISASVTEPLEHTAGEAMPPVEVTEK 532
QY 920 DIIAETPVLQTLPKGDADHDWVTSEVDFTSSEAVTATSTSEALRTEETVTEASGABETT 979
DB 533 DIIAETPVLQTLPKGDADHDWVTSEVDFTSSEAVTATSTSEALRTEETVTEASGABETT 592
QY 980 DMVSQVQLTDSPTTTEATPVQEVESGVLDTEEEERQTQAILQAVADKVKESQVPA 1039
DB 593 DMVSQVQLTDSPTTTEATPVQEVESGVLDTEEEERQTQAILQAVADKVKESQVPA 652
QY 1040 TVQRTGSKALEKVEEVEEDSEVLASEKEKDMPKGPVQVBAEHLAOGSETGQATPESLE 1099
DB 653 TVQRTGSKALEKVEEVEEDSEVLASEKEKDMPKGPVQVBAEHLAOGSETGQATPESLE 712
QY 1100 VPEVTADVDHVATCOVTKLQQLMEQAVAPESSETLTDSETNGSTPLADSTADTQODET 1159
DB 713 VPEVTADVDHVATCOVTKLQQLMEQAVAPESSETLTDSETNGSTPLADSTADTQODET 772
QY 1160 IDSQDSKATAAARQSQVTEEEAATAQKEEPTLFPNNVPAQEEHGEERGRDVLPTQOELT 1219
DB 773 IDSQDSKATAAARQSQVTEEEAATAQKEEPTLFPNNVPAQEEHGEERGRDVLPTQOELT 832
QY 1220 AAAPVLAKTEVQGEVDWLDGKVKKEQEVFVHSGPNSQKAADVTYDSEVMGVACQ 1279
DB 833 AAAPVLAKTEVQGEVDWLDGKVKKEQEVFVHSGPNSQKAADVTYDSEVMGVACQ 892
QY 1280 KESTEVQSLSEEGEMETDVEKRETKPEQVSEEGEQETAPEHEHNYGKPVLTLDMP 1339
DB 893 KESTEVQSLSEEGEMETDVEKRETKPEQVSEEGEQETAPEHEHNYGKPVLTLDMP 952
QY 1340 SERGKALGSLGSPSLPDQKAGCIEVQVQSLDTTQTAEAVEKVIETVVISSETGSPE 1399
DB 953 SERGKALGSLGSPSLPDQKAGCIEVQVQSLDTTQTAEAVEKVIETVVISSETGSPE 1012
QY 1400 CVGAHLLPAEKSSATGHWTLQHAEDTVPLGSPQAESEIPIIVTPAPESTLHPDLQGEIS 1459
DB 1013 CVGAHLLPAEKSSATGHWTLQHAEDTVPLGSPQAESEIPIIVTPAPESTLHPDLQGEIS 1072
QY 1460 ASQRESEEDKDPADGKSTAEIKVLAPEPELELESKSNKIVLNVITAVDQFA 1519
DB 1073 ASQRESEEDKDPADGKSTAEIKVLAPEPELELESKSNKIVLNVITAVDQFA 1132
QY 1520 RTETAPETHAYDSQTVPAACRLDSREPNCWTMKAKMKHPVQPPREDLQVLTVLEAW 1578
DB 1133 RTETAPETHAYDSQTVPAACRLDSREPNCWTMKAKMKHPVQPPREDLQVLTVLEAW 1191

RESULT 7
US-08-978-277A-2
; Sequence 2, Application US/08978277A
; GENERAL INFORMATION:
; APPLICANT: Gelman, Irwin H.
; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10112-0228
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/978,277A
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:

```

APPLICATION NUMBER: 08/665,401
FILING DATE: 18-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Richard S
REGISTRATION NUMBER: 26,154
REFERENCE/DOCKET NUMBER: A30558 - 165/34008
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-408-2558
TELEFAX: 212-765-2519
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1346 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
US-08-978-277A-2

Query Match      67.2%  Score 5421.5; DB 13; Length 1346;
Best Local Similarity 91.6%; Pred. No. 4.9e-304;
Matches 1098; Conservative 24; Mismatches 62; Indels 15; Gaps 7;

QY 387 MEAHQEVAAEVHVSIVTEKTEEEGGGGAEGGVVEGTGSLPPEKLAEPQVEAEBA 446
DB 1 MEAHQEVAAEVHVSIVTEKTEEEGGGGAEGGVVEGTGSLPPEKLAEPQVEAEBA 60
QY 447 BELMKSREMCVSGGDHQTLDLSPEEKLTPKHPREGIVSEVEMLSQERIKVQGSPLKCLF 506
DB 61 BELMKSREMCVSGGDHQTLDLSPEEKLTPKHPREGIVSEVEMLSQERIKVQGSPLKCLF 120
QY 507 SSSGKXKLSGKKOKGKRGGGDEEPGEYOHITHESPESADQKSSASPEPEETCTL 566
DB 121 SSSGKXKLSGKKOKGKRGGGDEEPGEYOHITHESPESADQKSSASPEPEETCTL 180
QY 567 EKGPLEAPQDGEAEGETTSDGKKRE---GITPMASFKNVTPKKRVRRPESDKEBEL 622
DB 181 EKGPLEAPQDGEAEGETTSDGKKRE---GITPMASFKNVTPKKRVRRPESDKEBEL 234
QY 623 EKVKASATLSTDSIVSEMODEVYKTVGEEBOKPEEPKRVDTSVSWELICVGSKKRARA 682
DB 235 EKVKASATLSTDSIVSEMODEVYKTVGEEBOKPEEPKRVDTSVSWELICVGSKKRARA 294
QY 683 SSSDDEGGRTLGSDSHRAEASKQKEAGTDVAPASTOEODQAQSSSPPEPAGSPSEGG 742
DB 295 SSSDIR-GRITLGGGOSORGGQORRRTDVPASTOEODQAQSSSPPEPAGSPSEGG 353
QY 743 VSTWSEFKRLVTPRKKSXSKLEEK-AEDSSVEQLSTELSPSREESVWS--IKKFIPIGRK 799
DB 354 VSTWSEFKRLVTPRKKSXSKLEEK-AEDSSVEQLSTELSPSREESVWS--IKKFIPIGRK 413
QY 800 KRAQKQEOATYEDSGPVEINEDDPNVPVPLSEINAVEREKMEQGTETLPOLLGANY 859
DB 414 KQOMROEOATYEDSGPVEINEDDPNVPVPLSEINAVEREKMEQGTETLPOLLGANY 472
QY 860 VSEELSKTLVHVVAVIDGTCAVTVSVEERSPSWISASVTEPLEHTAGAMPVEEVTEK 919
DB 473 VSEELSKTLVHVVAVIDGTCAVTVSVEERSPSWISASVTEPLEHTAGAMPVEEVTEK 532
QY 920 DIIAETPVLITLPECKADHDWMTSEVDFTSSEAVTATETSEALRTEVTEASGAETT 979
DB 533 DIIAETPVLITLPECKADHDWMTSEVDFTSSEAVTATETSEALRTEVTEASGAETT 592
QY 980 DMVSAVSQITDSDPTTEATPVQVEESGVLDTBEERQTOAILQAVADVKEESQVPATQ 1039
DB 593 DMVSAVSQITDSDPTTEATPVQVEESGVLDTBEERQTOAILQAVADVKEESQVPATQ 652
QY 1040 TYQRTGSKALEKEVEEBSEVLASEKEDVMPKGPVQVAGAEHLAQSEGTQATPESLE 1099

```

```

DB 653 TYQRTGSKALEKEVEEBSEVLASEKEDVMPKGPVQVAGAEHLAQSEGTQATPESLE 712
QY 1100 VPEVTAVDVHATQVVKLQOLMEQAVAPSSSETLTSTENGSLPLADSDTADTQODET 1159
DB 713 VPEVTAVDVHATQVVKLQOLMEQAVAPSSSETLTSTENGSLPLADSDTADTQODET 772
QY 1160 IDSQSKATAAVROSQVTEEBAAFOKEEPTLPPNNPAOEHEGEGRPDLPEPTQOELT 1219
DB 773 IDSQSKATAAVROSQVTEEBAAFOKEEPTLPPNNPAOEHEGEGRPDLPEPTQOELT 832
QY 1220 AAAPVPAKTEVGEQGEVDMLDGEKVEEDEVFVHSGPNSOKADVTVTDSRWAGCOE 1279
DB 833 AAAPVPAKTEVGEQGEVDMLDGEKVEEDEVFVHSGPNSOKADVTVTDSRWAGCOE 892
QY 1280 KESTEVQSLSEEBEMETDVKEKRETKPBOVSSEGOETAPAPHEGTGKPVLTLMPS 1339
DB 893 KESTEVQSLSEEBEMETDVKEKRETKPBOVSSEGOETAPAPHEGTGKPVLTLMPS 952
QY 1340 SERGKALGSLGSPSLPDQKAGCIEVQVOSLDTTVQTAEAVERKVIETVVISETGESPE 1399
DB 953 SERGKALGSLGSPSLPDQKAGCIEVQVOSLDTTVQTAEAVERKVIETVVISETGESPE 1012
QY 1400 CVGAHLPAEKSSATGHWTLQHAEDTVPLGPESQASIPITVPAPBSTLHPDLQSEIS 1459
DB 1013 CVGAHLPAEKSSATGHWTLQHAEDTVPLGPESQASIPITVPAPBSTLHPDLQSEIS 1072
QY 1460 ASQERSEEDKPPDAGPDADKESTAIKVLKAPPEILESKNKVLYLVITQAVDOFA 1519
DB 1073 ASQERSEEDKPPDAGPDADKESTAIKVLKAPPEILESKNKVLYLVITQAVDOFA 1132
QY 1520 RTEPAETHAVDSQTOVPACRLDSREPNCWTXKXDKMKHPVOPREDLOVLTVLEAM 1578
DB 1133 RTEPAETHAVDSQTOVPACRLDSREPNCWTXKXDKMKHPVOPREDLOVLTVLEAM 1191

RESULT 8
US-09-902-432-2
; Sequence 2, Application US/09902432
; GENERAL INFORMATION:
; APPLICANT: Irwin H. Gelman
; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
; FILE REFERENCE: A30558-A-FWC-A 070156, 0597
; CURRENT APPLICATION NUMBER: US/09/902,432
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 08/978,277
; PRIOR FILING DATE: 1997-11-25
; PRIOR APPLICATION NUMBER: 08/665,401
; PRIOR FILING DATE: 1996-06-18
; PRIOR APPLICATION NUMBER: 08/635,121
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1346
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-902-432-2

Query Match      67.2%  Score 5421.5; DB 23; Length 1346;
Best Local Similarity 91.6%; Pred. No. 4.9e-304;
Matches 1098; Conservative 24; Mismatches 62; Indels 15; Gaps 7;

QY 387 MEAHQEVAAEVHVSIVTEKTEEEGGGGAEGGVVEGTGSLPPEKLAEPQVEAEBA 446
DB 1 MEAHQEVAAEVHVSIVTEKTEEEGGGGAEGGVVEGTGSLPPEKLAEPQVEAEBA 60
QY 447 BELMKSREMCVSGGDHQTLDLSPEEKLTPKHPREGIVSEVEMLSQERIKVQGSPLKCLF 506
DB 61 BELMKSREMCVSGGDHQTLDLSPEEKLTPKHPREGIVSEVEMLSQERIKVQGSPLKCLF 120
QY 507 SSSGKXKLSGKKOKGKRGGGDEEPGEYOHITHESPESADQKSSASPEPEETCTL 566

```

```
Db 121 SSSGLKLLSGKKKGKGGGDEBPGEYQVHIHTESPESADEQKGESSASSPEPEETTCL 180
Qy 567 EKGLEAPQGEABEGTTSDEKKE---GITWAFSKKMTPKKVRPSPESDKKEEL 622
Db 181 EKGLEAPRMGKURKELLR-GEKKDHSLG-----LQKDGDTQETVRRPESDKKEEL 234
Qy 623 EKVSATLSSTDSTVSEWQDEVKTVEGEQPEBKRRVDTSVSWEALICVSSSKRKARKA 682
Db 235 EKVSATLSSTDSTVSEWQDEVKTVEGEQPEBKRRVDTSVSWEALICVSSSKRKARKA 294
Qy 683 SSSDEGGPRTLGGDHSRAEASAKDEAGTDVAPASTQEQDQAQSSSPAGSPSSEGG 742
Db 295 SSSDIR-GRPTLGGGQSQRGQRORSTRDAPASTQEQDQAQSSSPAGSPSSEGG 353
Qy 743 VSTWESPKRLVTPRKSKSKLEEK-ADSSVEQLSTIEBPSRESWVS--IKFIPGRRK 799
Db 354 VSTWESPKRLVTPRKSKSKLEEKAGRTLVGAGCPLRSNRVEKNLGFPLRSSPDGGR 413
Qy 800 KRADGKEQATVEDSGPVEINEDPNVPVAVVPLSEYNAREKMEAGNTELPOLLGAVY 859
Db 414 KQMGREQATVEDSGPVEINEDPDVAVVPLSEYDAVEREKMEAGNAELPSCWGCV- 472
Qy 860 VSELSKTLVHTVSVAVIDGTRAVTSVEERSPSWISASVTEPLEHTAGEAMPVVEYTEK 919
Db 473 VSELSKTLVHTVSVAVIDGTRAVTSVEERSPSWISASVTEPLEHTAGEAMPVVEYTEK 532
Qy 920 DIIAETPVLVTQTLPEGKAHDDMVTSEVDFTSEAVTATITSEALRTEEVTEASGABETT 979
Db 533 DIIAETPVLVTQTLPEGKAHDDMVTSEVDFTSEAVTATITSEALRTEEVTEASGABETT 592
Qy 980 DMYSAVSOLTDSPTTTEATPVQVEGVLDTBEEROTCAILQAVADVKKESSQVPATQ 1039
Db 593 DMYSAVSOLTDSPTTTEATPVQVEGVLDTBEEROTCAILQAVADVKKESSQVPATQ 652
Qy 1040 TVQRTGSKALEKVEVEEDEDSEVLASEKEKVMKPGVQVQEAHAELAGCSETGQATPSLE 1099
Db 653 TVQRTGSKALEKVEVEEDEDSEVLASEKEKVMKPGVQVQEAHAELAGCSETGQATPSLE 712
Qy 1100 VPEVTADVHVATCOVLIKLOLMQEAQVAPESSETLTDSETNGSTPLADSDTADTQODET 1159
Db 713 VPEVTADVHVATCOVLIKLOLMQEAQVAPESSETLTDSETNGSTPLADSDTADTQODET 772
Qy 1160 IDSODSKATAVRQSQVTEBEANATAQKEEPTLBNPNVPAQEEHGEERGRDVLPTQOELT 1219
Db 773 IDSODSKATAVRQSQVTEBEANATAQKEEPTLBNPNVPAQEEHGEERGRDVLPTQOELA 832
Qy 1220 AAAPVPLAKTEVGOGEVDWLDGKVKKEQEVFVHSGPNSQKAADVTDSEVMGVACQE 1279
Db 833 AAAPVPMQKTEVGOGEVDWLDGKVKKEQEVFVHSGPNSQKAADVTDSEVMGVACQE 892
Qy 1280 KESTEVQSLSEEGEMETDVEKRETKPEQVSEEGEQETAAPHEGTYGKPVLTLDMP5 1339
Db 893 KESTEVQSLSEEGEMETDVEKRETKPEQVSEEGEQETAAPHEHNYGKPVLTLDMP5 952
Qy 1340 SERKALGSLGGSPSLPDQKAGCIEVQVQSLDITVTQTAAVAKVJETVVISETGSPE 1399
Db 953 SERKALGSLGGSPSLPDQKAGCIEVQVQSLDITVTQTAAVAKVJETVVISETGSPE 1012
Qy 1400 CVGAHLLPAEKSSATGHWTLQHAEDTVPLGPESQAESIPIITVPAPESTLHPDQGEIS 1459
Db 1013 CVGAHLLPAEKSSATGHWTLQHAEDTVPLGPESQAESIPIITVPAPESTLHPDQGEIS 1072
Qy 1460 ASQRESEEDKPDAGPADGKESTAEIKVLKAEPEILELESKNKVLNVIVQTAVDQFA 1519
Db 1073 ASQRESEEDKPDAGPADGKESTAEIKVLKAEPEILELESKNKVLNVIVQTAVDQFA 1132
Qy 1520 RTETAPETHAYDSQTPACRLDSREPNRCWTMKDKMKHPVPPQPPREDIQLVTLBAW 1578
Db 1133 RTETAPETHAYDSQTPAMRLDSREPNRCWTMKDKMKHPVPPQPPREDIQLVTLBAW 1191
```

RESULT 9

US-60-389-987-2097

```
; Sequence 2097, Application US/60389987
; GENERAL INFORMATION:
```

```
; APPLICANT: Ghosh, Soumitra S.
```

```
; APPLICANT: Fahy, Eoin D.
```

```
; APPLICANT: Zhang, Bing
```

```
; APPLICANT: Gibson, Bradford W.
```

```
; APPLICANT: Taylor, Steven W.
```

```
; APPLICANT: Glenn, Gary M.
```

```
; APPLICANT: Warnock, Dale E.
```

```
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
```

```
; FILE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
```

```
; FILE REFERENCE: 660088.465P2
```

```
; CURRENT APPLICATION NUMBER: US/60/389,987
```

```
; CURRENT FILING DATE: 2002-06-17
```

```
; NUMBER OF SEQ ID NOS: 3025
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 2097
```

```
; LENGTH: 1782
```

```
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
```

```
; US-60-389-987-2097
```

```
Query Match
```

```
Best Local Similarity 53.5%; Score 4318.5; DB 27; Length 1782;
```

```
Matches 961; Conservative 189; Mismatches 398; Indels 107; Gaps 34;
```

```
Qy 1 MGAGSSTEQRSPQ-PAGSDTPSELVLSHGHPAAEAS-GAAGDPADA--DPATKLPQKNG 56
```

```
Db 1 MGAGSSTEQRSPQPEGSSTPAEPSPGGPSAAAPDTPADPAIAASDPATKLLQKNG 60
```

```
Qy 57 QLSSVNGVAFQGVHVHVENQEQ-----EEVVDEVDGQRESDEVREKDR 102
```

```
Db 61 QLSTINGVAFQDELSQLQEGDLNGKQALNGQAGALNSQEEVEVITVEVQGRSDSDVSKRDS 120
```

```
Qy 103 VEEMANSTAVEDITKDGQEBTSEITQIPASENNVVEVMQPAESQANDVGPKVFKFVG 162
```

```
Db 121 DKEMATKSAVVHDDITDQGEETPEITQIPSSSESNLEUTQPTESQANDIGFKKVFQVG 180
```

```
Qy 163 FKFTVKDKNEKSDTVQLLTVKKDEGEAEASVAGADHQBSPSVETAVGESAKSESELKQS 222
```

```
Db 181 FKFTVKDKTEKPDVTQLLTVKKDEGEA--AGAGDHKDPSL--GAGEAASESEPKQS 235
```

```
Qy 223 TEKQEGTLKQEQSSTIPIQAESDQAAEBEAKDEGEKEKEPTKSPESPSPVNETTS 282
```

```
Db 236 TEKPEETLKREQSHABISPPAESQAV-BECKEGBEKEKEPKSKAESPTSPVTSETGS 294
```

```
Qy 283 SPKKFTTHGWAGWRKKTSPKKSKEDDLTAETAEKKEQEAEEKVDEEKEKTEPASEE----- 337
```

```
Db 295 TFKKFTQGWAGWRKKTSPKPKDEDEVEASEKKKEQEPKVDTEEDGKAEVASEKLTASE 354
```

```
Qy 338 ----QBPEDTDQARLSADYKVELPLEDQVGDLEASSBEKCAPLATEVDFDEKWEAHQ-E 392
```

```
Db 355 QAHPQEPASAEHPRLISAIEYKVELPSEBQVSGSQSPSEKAPLATEVDFDEKLVHQB 414
```

```
Qy 393 VVAEVHVTVEKTEBQGGGGAEGGVVVEGTGESLPPKLAEPQEVQPAEAPABLMS 452
```

```
Db 415 VVAEVHVTVEERTSEQ-----KTEVEETAGSVPASELVEMDAEPQEAEPKSLVKL 466
```

```
Qy 453 REMCVSGGDBHTQTLTDLSPBEKTLPKHPEGIIVSEVEMLSQERIKVQGSPLKCLFSSSGLK 512
```

```
Db 467 KETCVSGEDPTQAGADLSKPEKVLSPKPPGVVSEVEMLSQERMKVQGSPLKCLFTSTGLK 526
```

```
Qy 513 KLSGKKQKGRGGGDEBPGEYQVHIHTESPESADEQKGESSASSPEPEETTCLGKPLE 572
```

```
Db 527 KLSGKKQKGR--GGGDESGEHTQVPADSPDSQEEQKGESSASSPEPEITCLEKGLAE 595
```

```
Qy 573 APODGAEBEGTTSDEGKKEGITPWASFKNVTPKRVRRPSPDSKKEBLEKYKSLTSS 632
```

```
Db 586 VQODGAEBEGATSDGKKREGVTPWASFKNVTPKRVRRPSPDSKDEBLDKVKSLTSS 645
```

```
Qy 633 TDSVTSEMQDEVKTVEGEQPEBKRRVDTSVSWEALICVSSSKRKARKASSDDGGR 692
```

```
Db 646 TESTASEMQEEMKGSVEEPKPKPKKVDTSVSWEALICVSSSKRKARRGSSSDEEGPK 705
```

```

QY 693 TLGDSHRAEASKDKKXGTDVAVPASTQDQOAGSSSPREPAGSPSEGBGVSTWMSFKRL 752
Db 706 AMGGHQKADKXGKXGTDGDIILAGSQHDPGQSSSSPEQSGPTSEGBGVSTWMSFKRL 765
QY 753 VTPRKSKSKLEEKAD-----SSVEQLSTEIPRSREESVSIKKFIPGRKKRAAGKQBO 808
Db 766 VTPRKSKSKLEEKEDSLAGSGVSHSTPDTIPGKEBSVSIKKFIPGRKKRPPGKQBO 825
QY 809 ATVEDSGEIVINEDDPNVPVAVPLSEYNAREKME---AQNTLPLQLLGAIVVSELS 865
Db 826 APVEDAGETGANNEDSDVAVVPLSEYDAVEREKMEAOQAQASAPQEQKATEVSKELS 885
QY 866 KTLVHTSVAVVIDGTRAVTSVEERSPSWISASVTEPLEHTTAGEAMPVHEVTEKILIA-E 924
Db 886 ESQVHMMMAAAVADGTRAAITIIERSPSWISASVTEPLEVEAEALLTVEVIEREVIABE 945
QY 925 ETPVLTOTLPECKADHDMVNTSEVDPTSEAVATETSEALRTVEVYASGAETTDWISA 984
Db 946 EPTVTEPLEPENREARQDVTVSEAEITTPRAVTAETPAGPLGAEGETMSAAETTEMVISA 1005
QY 965 VSQLTDSPTTEATPVQEVESGVLDTEEEERQTAIIQAVADKVKESQVPAE----- 1038
Db 1006 VSQLTDSPTTEATPVQEVESGVLDTEEEERQTAIIQAVADKVKESQVLPOTGQBEDV 1065
QY 1039 -QTVQPTGSKALEKVBVEVDESEVLASEKQVMKGPVQEGAEHLAOGSETQATPES 1097
Db 1066 LQPVORA-----EAEPRPEQAEASGLKKEQTVVLKVDQAEKTEPTFGKVVQQTTPES 1119
QY 1098 LE-VREVTAADVH---VATQCV-----IKLOOL-MEQAVAPASSELTLDSETNGSTPLAD 1147
Db 1120 FKAQOVVESISSLVLTQAEHTLAGVKSQEWMEQALPPSVETPTDSEIDGSTFPVAD 1179
QY 1148 SPTADGTOQDETITDSQSKATAVAVRQSVYTEEBAATQKEEPTLPNNVPAQEHGEERG 1207
Db 1180 FPARPTQODELVEIHENBVAISGTSQGTAEAVPAQKERPRA-PSSVFPQETKEQSK 1238
QY 1208 -RDVLEPTQOELTAAVAVPLATEVQEGEVDMLDGEKXK-----EQQVFNHSG--PMS 1259
Db 1239 MEDTLEHTDKEVSVEVTSIISTKEGTQ--EADQYADEKTKDVPFPGLEGSIDTGIIVSR 1296
QY 1260 QKAADVTVSEVMGVAGCOEKSESTEVQ-----LSLEEGEMETDVKERETKPEQVSE 1314
Db 1297 EKVTEVAKGEGSTEAECKKDALDELQSHAKSPSPSVBEREMVAVYERETKTEAPTHVNEE 1356
QY 1315 G-EOETAAEHEGTYGKPVLTLDMPSEERKALGSLGS-PSLPDQKAGCIEVQYOSLD 1372
Db 1357 KLEHETAVVSEVSKQLQTVNVPIIDAKAVSISLGSPPCLQGEAEAVCTKIQVQSSB 1416
QY 1373 TTVTQTAAEVAVKVI--ETVVISETGSPRCVGAHLLPAKSSATGHWTLQHAEDTVPLG 1430
Db 1417 ASFTLTAAAEKKEVLEGTANILLETGTELPAGAHVLEKSSSEKHNEDFAHNGEDAVPTG 1476
QY 1431 PESQAEISTITVTPABESTLHPDLOGEIASQRESEEDDKPADGDAQKSTALEKXL 1490
Db 1477 PQCQAKSTVIVSATTKGGLSSDLSEKETSLSKWSDEVDEQVACQEV--KISVALIEDU- 1533
QY 1491 KAEPE--IILELSSKNKIVLVANVQTAVDQFART-ETAPETHAVDSQTVPAKCLDSREBN 1547
Db 1534 --EPENGILIELETKSKVLQVNIITQAVDQFVTRTEFTATEMLTSELQTHAVIKADQDAG 1591
QY 1548 KCMWK-----MDAKMKHPVQFREDLQVLTVLEA 1577
Db 1592 QETKEGEEPPQASQODETPTISAKESSESTAVQQA 1626

```

```

: APPLICANT: Gibson, Bradford W.
: APPLICANT: Taylor, Steven W.
: APPLICANT: Glenn, Gary M.
: APPLICANT: Wainock, Dale E.
: TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
: FILE REFERENCE: 660088.465P2
: CURRENT APPLICATION NUMBER: US/60/412.418
: CURRENT FILING DATE: 2002-09-20
: NUMBER OF SEQ ID NOS: 3025
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 2097
: LENGTH: 1782
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-60-412-2097

Query Match 53.5%; Score 4318.5; DB 27; Length 1782;
Best Local Similarity 58.1%; Pred. No. 4.7e-240;
Matches 961; Conservative 189; Mismatches 398; Indels 107; Gaps 34;

QY 1 MGAGSTEQRSPEQ-PAGSDTPESELVLSGHCPAAEAS-GAAGDPADA--DPATYLPQKNG 56
Db 1 MGAGSTEQRSPEQPEPESSTPAPEPESGGGSAEAPDTPADPAIAASDPATKLLQKNG 60
QY 57 QLSVNGVABEGDVHVGNEQO-----EEEVNVEDVQORESEDEREKDR 102
Db 61 QLSVINGABEDELISLQGDNDNGKALNGQALNSQEEBIVITEVGQRSEBVSRKDS 120
QY 103 VEEMAASTAVATEDITKQOEETSEIIQIPASENNVBERMVPAASQANDVGFKIVFKVG 162
Db 121 DKMAATKAAVVDITDQOEETPEIIRIIPSESNLELTQPTSQANDIGFKIVFKVG 180
QY 163 FKFTYKDKNEKSDTVOLLTVKQDEBGAASVQAGDHQEPSTVAVGASAKSELSKOS 222
Db 181 FKFTYKDKNEKSDTVOLLTVKQDEBGA--AGAGDHKDPSTL--GAGEAASKSESPKOS 235
QY 223 TEKQOGLTKQOSSTEIPLQAESDQAAEEAKDGESEKQEKPEKSPSPSPVNSSTTS 282
Db 236 TEKPEETIKRQSHAEISPPAESQAV--EECKEGESEKQEKPSKASPTSPVTSSTGS 294
QY 283 SPKKEFTGMAQMKRKTSPFKSKKEDLETAERKKEQAEKVEDEKTEPEASEE----- 337
Db 295 TYKKEFTGMAQMKRKTSPFKPKEDVYASERKKEQOEPEKKTDEBDGAEVASEKLTSE 354
QY 338 ----QEPREDTQARLSADYKVELPLEDQYDLEASSESEKCAPLATEVPDEKMEANO-E 392
Db 355 QAHPEPAPESAHBRLSAEYKVELPSEBOVSGSGPSEKAPLATEVPDEKIEVHQEE 414
QY 393 VVAEVHNVSTVEKTEBEQGGGABEGVNVVEGTGSELPPEKLAEPQEVQAEAPAEELMKS 452
Db 415 VVAEVHNVSTVEERTEEQ-----KTEVEETAGSVPAEELVENDAEQAEAPAEELVKL 466
QY 453 REMCVSGGDHQLTMDLSPDEKTLPRKHPEGIVSEVEMLSQERIKYQSGPLKTLFSSGLX 512
Db 467 KETCVSGSDPLQAGDLSPDEKVLSPPEGVISSEVEMLSQERIMKYQSGPLKTLFSTGLX 526
QY 513 KLSGKKQKGGKGGGDEEPGEVQHITHSPSADBDQKGESSASAPSEDETTCLEKGPL 572
Db 527 KLSGKKQKGRK-GGDEBSGHTQVPADSPSOERQKGESSASAPSEDETTCLEKGLAE 585
QY 573 APQDEABEAGTTSQDEKRGREGITTPWASFKXVNTPKKRRRRSESEKLELKVYSATLSS 632
Db 586 VOQDEABEAGTSDDEKREBGTTPWASFKXVNTPKKRRRRSESEKLELKVYSATLSS 645
QY 633 TDSVSEMODEVKTIVGEEQKPEEPKRRVDTGSVSWALICVSSSKKRRARAKASSDDEGPR 692
Db 646 TBSVSEMODEVKTIVGEEQKPEEPKRRVDTGSVSWALICVSSSKKRRARAKASSDDEGPR 705
QY 693 TLGDSHRAEASKDKKXGTDVAVPASTQDQOAGSSSPREPAGSPSEGBGVSTWMSFKRL 752
Db 706 AMGGHQKADKXGKXGTDGDIILAGSQHDPGQSSSSPEQSGPTSEGBGVSTWMSFKRL 765

```

```
Qy 753 VTPRKSKSLEKAEAD-----SSVEQLSTETEPSREESWVSIKKFIPGRRKKRADGKQEQ 808
Db 766 VTPRKSKSLEKSESDSIAGSGVHSTPDTEPGKEESWVSIKKFIPGRRKKRPDQKQEQ 825
Qy 809 ATVEDSGVEINEDDPNVPAVPLSEYNAREKME---ACNTELPOLLGAVVYSBELS 865
Db 826 APVEDAGPTGANEDSDVPAVPLSEYDAVERKWEAQAKSAEQPEKKAATEVSKELS 885
Qy 866 KTLVHTVSVAVIDGTRAVTSVEERSPSWISASVTEPLEBHTAGEAMPVVEEVTEKDI1A-E 924
Db 886 ESQVHMAAAVADGTRAAITIEERSPSWISASVTEPLEQVEAEAAALLTEELEREVIAEE 945
Qy 925 ETPVLTOPLPGKAHDDMTVSEVDFTSEAVTATETSEALTETSEALTETSEALTETSEALT 984
Db 946 EPPVTEPLPENRARGTGVVSEAELETPVAAETAETAGLGAEBEGTEASAAEETEMVSA 1005
Qy 985 VSQLTSDPTTEETATPVQEVESGLVDTTEEBERQTAIIQAVADVKKEESQVPAT----- 1038
Db 1006 VSQLTSDPTTEETATPVQEVESGLVDTTEEBERQTAIIQAVADVKKEESQVPAT----- 1065
Qy 1039 -QTVORTGSKALEKVEEVEEVEEVEEVEEVEEVEEVEEVEEVEEVEEVEEVEEVEEVEE 1097
Db 1066 LQPVQRA-----EAERPEEQAEASGLKEDTVLVKDAQEAQTEPFTQGVVGGTTPE 1119
Qy 1098 LE-VPEVTADVDH---VATCOV-----IKLOOL-MEQAVAPESSETLTDSSETNSTPLAD 1147
Db 1120 FEKAPQVTEESSELVTTQAEITAGLAGVQSEMYMEQAIPDPSVETPTDSDSETGSTPVAD 1179
Qy 1148 SDAADGTQODTETIDQSKATAAARQSOVTEEEAATAQKEPSTLPNNVPAQEBEGREP 1207
Db 1180 FDACTTQKDIIVEIHENEVASGTQSGGTEAEAVPAQKERRPPA-PGSFVQEETKQSK 1238
Qy 1208 -RDVLEPTQBELTAAAVPLAKTEVQGEVDWLDGEKVK-----EEOEVFVHSG---PNS 1259
Db 1239 MEDTLEHTKEVSVETVSIKTEGTQ--EADQVADKTKDVPPFEGLEGSIDGTIVSR 1296
Qy 1260 QKAADVITYDSEVMGACQEKESTEVQS-----LSLEGEMETDVEKREKTKTEQVSE 1314
Db 1297 EKVTEVALKGBGTBEAECKODALELOSHAKSPSPVEREMVQVVEREKTAEAEPTHVNEE 1356
Qy 1315 G-EQETAAPHEGTVGKPLVLTDMPPSSRKGALGSLGGS-PSLPDQDKAGCIEVQVOSLD 1372
Db 1357 KLEHETAVTVSEEVSKQLQTVNVPIIDGAKESVSLGSPPCLGQEEAVCTKIQVQSSE 1416
Qy 1373 TTVTQTAABAEKVI--ETVWISSETGESPECVGAHLLPAEKSSATGGHWTLOHAEDTVP 1430
Db 1417 ASFTLTAABAEKVLGETANILETGETLEPAGAHVLEEKSEKNEKEDFAAHPGEDAVPTG 1476
Qy 1431 PESQAESPIITVPAPESTLHPDIOGEISASQERSBEEDKPDGADGKESTAIKVL 1490
Db 1477 PDCQAKSTPVIVSATTCKGLSSDLEGEKTTSLKWKSDSEVDQVACQEV--KVSVAIEDL- 1533
Qy 1491 KAEP--ILELESKNKIVLNVIOQTAVDQPART--ETAPETHAYDSQTVPAACRLDSREP 1547
Db 1534 --EPENGILELETSSKLVQNIITQAVDQFVRTTETATEMTLSETLQTAHVIKADSDQAG 1591
Qy 1548 RCWTK-----MKDAMKHPVPQPPREDLQVLTVLEA 1577
Db 1592 QETEKEGEPQASQADEPITSAKESESTAVGQA 1626
```

RESULT 11

```
US-09-738-877-3
; Sequence 3, Application US/09738877
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan
; APPLICANT: Weils, Stephen J.
; APPLICANT: Glynn, Richard
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF ANGIOGENESIS, COMPOSITIONS, AND MET
; TITLE OF INVENTION: SCREENING FOR ANGIOGENESIS MODULATORS
; FILE REFERENCE: A-69806/DJB/JJD
; CURRENT APPLICATION NUMBER: US/09/738,877
```

```
; CURRENT FILING DATE: 2000-12-15
; PRIOR APPLICATION NUMBER: US 09/637,977
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: US 60/148,425
; PRIOR FILING DATE: 1999-08-11
; PRIOR APPLICATION NUMBER: PCT/US 00/22061
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 3
; LENGTH: 1781
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-738-877-3
```

```
Query Match 53.2%; Score 4294; DB 21; Length 1781;
Best Local Similarity 57.9%; Pred. No. 1.2e-238;
Matches 959; Conservative 187; Mismatches 401; Indels 108; Gaps 35;
```

```
Qy 1 MGAGSSTEQRSRPGQ-PAGSDTPSELVLSHGHPAAEAS-GAAGDPADA--DPATKLPQKNG 56
Db 1 MGAGSSTEQRSRPGQSGSTPAEPSPGGPSAEAPDTTADPAIAASDPATKLLQKNG 60
Qy 57 QLSVNGVAGSDVHVQENQEQ-----EEVVDEDVQRESESDVREKOR 102
Db 61 QLSSTINGVAGQDLSLQEGDLNGKQALNGQGALNSQEEEVIVTEVQGRDSEDVSRDS 120
Qy 103 VEEAANSTAVEDITKQGEETSEIIQIIPASNNVEMVQPAESQANDVGFKKVFPVFG 162
Db 121 DKEMATKSAVVHIDTDDQENRN-IEQIPSSSNLEELTQPTESQANDIGFKVFPVFG 179
Qy 163 FKFTVKDKNEKSDTVOLLTVKKDEGEBAEASVAGAGHQBPSPVETAVGESAKSELKQS 222
Db 180 FKFTVKDKTEKPTVQLLVKKDEGEBA--AGAGDHQDPSL--GAGEAASKESEPKQS 234
Qy 223 TEQOEGTLKQEQSTETPLQAESDQAAEBAKQEBKQEBKPTKSPESPSPVNSSTTS 282
Db 235 TEKPEETLKREQSHAETSPPAESQAV-ECKEKEGEEKQEKPSKSAESPTSPVTSSTGS 293
Qy 283 SFKFFFTGHWAGWRKTSFKKSKEDDLETAERKEQAEKVEEKEKTEPASEE----- 337
Db 294 TFKKFTTQGWAGWRKTSFKPKEDVEASEKKQEBQPEKVDTEEDGKAIVASEKLTASE 353
Qy 338 ----QEPADTDQARLSADYKVELPLEDQVGLDLEASSEEKCAPLATEVEFDEKMEAHQ-E 392
Db 354 QAHPQEPAESAEHPRLSAEYKVELPSEEQVSGSQGPSEKPAFLATEVEFDEKIEVHQEE 413
Qy 393 VVAEVHVSTVEKTEEBQGGGGBAGGVVVEGTGESLPPEKLABPQEVPOBAEAEELMKS 452
Db 414 VVAEVHVSTVEERTTEQ-----KTEVEETAGSVPAEELVGMADAPQBAEPAKELVKL 465
Qy 453 REMCVSGDDHTQLTDLSPBEKTLPKHPEGIVSEVEMLSOERI KVOGSLPKLKLFPSSGLK 512
Db 466 KETCVSGEDTQAGADSLPDEKLSKPEGVVSEVEMLSOERMKVOGSLPKLKLFTSTGLK 525
Qy 513 KLSGKKQKGRGGRGDEEPEGEYQIHTESPEADEQKGESSASSPEPEITTCLEKGPLE 572
Db 526 KLSGKKQKGR--GGGDEESGEHTQVPADSPDSQEQKGESSASSPEPEITTCLEKGLAE 584
Qy 573 APQDGAEBEGTSDGKKBKREGITPWASFKQWTPKRVRRPSPESDKEEELKVKGSATLSS 632
Db 585 VQDQGEAEAGATSDGKKBKREGVTPWASFKQWTPKRVRRPSPESDKEDELKVKGSATLSS 644
Qy 633 TDSITVSEMDEKTVGEEQKPEEKRRVDTSVSWEALI CVGSSKKKARKASSSDDEGPR 692
Db 645 TESTATSEMDEKTVGEEQKPEEKRRVDTSVSWEALI CVGSSKKKARKARRSSDEEGPK 704
Qy 693 TLGDSHRAEEASKDKEAGTDVAPASTQEODAQGSSSPPEAGSPSEGEVSTWESFKRL 752
Db 705 AMGDGHQKADGAKDXTGDTGLAGSQEHDPQGGSSSPQAGSPTEGEVSTWESFKRL 764
Qy 753 VTPRKSKSLEKAEAD-----SSVEQLSTETEPSREESWVSIKKFIPGRRKKRADGKQEQ 808
```

```
Db 765 VTPRKSKSLEEKSESDSIAGSVESHSTPDTPEPKGEESWISIKFIPGRKKRPDQKQO 824
Qy 809 ATVEDSGVEINEDDPNVPVAVPLSEYNANVEREKME--AQNTLPLQILGAVVSEELS 865
Db 825 APVEDAGFTGANEDSDVPAVAVPLSEYDAVEREKMEAOQAQKGAQPEQKATEVSKELS 884
Qy 866 KTVHTVAVAVIDGRATVSVSESPSWISATVTEPLEHTAGEMAPVPEVTEKIIA-E 924
Db 885 ESQVMMMAAAGDGRAATIIIEKSPSWISATVTEPLEVEAEALLTVEVLEKREVIABE 944
Qy 925 ETPVLTQTLPECKADHDMVTSEVDFTSSEAVATETSEALRTVEEYVTEASGAETIDMVA 984
Db 945 EPPVTEPELPEKREARAGDTVVSSEALTPRAVTAETAGLGESEGTESAABETTEMVSA 1004
Qy 985 VSQLTSDSPDTEEATPVQVESGVLDTEEEERQTAIIQAVADKYKESQVPAE----- 1038
Db 1005 VSQLTSDSPDTEEATPVQVEGEGVPDIEQERRTQEVLAQVAEKYKESQVPGTGPEDV 1064
Qy 1039 -QTVQRTSKALEKVEVEDESEVLASEKEDVMKGPVQEGAGHLAGSEGTQATPES 1097
Db 1065 LQPVQRA-----EAEKREPEQAEASGLKKEEDVILKVAQEKTEPFTQKVGQITPES 1118
Qy 1098 LE-VREVTADVDH---VATQCV-----IKLQOL-MEOAVAPESSELTLPSETNGSTPLAD 1147
Db 1119 PEKAPQVTESSISELVTTCQAEITLAGVKSQEMWEOAIPDSVETPTDSETDSTPVAD 1178
Qy 1148 SPTADGTQODETIIIDSQSKATPAVRSQVTEBEAATAQKEEBSLTPNNVPAQEEHEBEG 1207
Db 1179 PDAPPTQODELVEIHEENEVAVSGTQSGTEAEVAPQAKERPRA-PSSEVFQEEETKEQSK 1237
Qy 1208 -RDVLEPTQOELTAAVAVPLATEVGOEGEVMLDGEKTK-----EEQVFNHSG--PNS 1259
Db 1238 MEDTLEHTDKEVSVETVSLSTKEGTQ--EADQVADKTKDVPFEGLEGSIDTGTIVSR 1295
Qy 1260 QKADAVTADSEWAGVAGCOEKESTEVOS-----LSLEGEMETDVEKEKREKPCOVSE 1314
Db 1286 EKVTEVALKGEETGEAECKKDALBELQSHAKSPSPVESEMVQVHERETEMEPHVNNE 1355
Qy 1315 G-EOETAAPEHGIYQKPYLTLDMPSSERKALGSLGGS-PSLPDQKAGCIEVOVQSID 1372
Db 1356 KLEHETAVTSEVSEVSKQLQTVNVPFIIDGAKVSLSESGPPCLQGEBAVCTKIYQSSSE 1415
Qy 1373 TTVVTOAETAEVKTI--ETVVISETGESPECVGAHLLPAKKSATGHTLQIAEDTVPILG 1430
Db 1416 ASFTTLTAAEEKVKVGETNIILETGETLEPAQAHVLEKSSSEKNEDEFAHPEGEAVPFG 1475
Qy 1431 PPSQAESEPIIYTPAESTLHPDLQGEISASQRESSEEDKPDAGDAGKESTALEKVL 1490
Db 1476 PDCQAKSTVIVSATTKKLSLSDLEGEKTSLSKMSDEVDEQVACOEV--KVSVAIEDL- 1532
Qy 1491 KNAPE--ILEESKSNKIVLAVIQTAVDQPART-ETAPETHAVDSQTVPAQRDLDSREBN 1547
Db 1533 --EPENGILELTKSSKLVQNTIIQTAVDQFVTEETATMLTSELQTOHVAIKADSQDAG 1590
Qy 1548 RCWTK----MKDAMKHPVPOPRDLOVLTYLEA 1577
Db 1591 CETEREGEERQASADEPTPITSAKESSESTAVGOA 1625

RESULT 12
US-09-961-403-13
; Sequence 13, Application US/09961403
; GENERAL INFORMATION:
; APPLICANT: HE-STUMPP, HOLGER
; APPLICANT: HAENDLER, BERNARD
; APPLICANT: KRAETZSCHMAR, JOERN
; APPLICANT: KREFT, BERTHOLT
; APPLICANT: WINTERHAGER, ELKE
; APPLICANT: REGIDOR, PEDRO
; APPLICANT: SCOTTI, SIMONE
; TITLE OF INVENTION: METHOD FOR IN VITRO DIAGNOSIS OF ENDOMETRIOSIS
; FILE REFERENCE: SCH-1789
; CURRENT APPLICATION NUMBER: US/09/961,403
```

```
; CURRENT FILING DATE: 2001-09-25
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 1781
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-961-403-13

Query Match          53.2%; Score 4294; DB 23; Length 1781;
Best Local Similarity 57.9%; Pred. No. 1,2e-238;
Matches 959; Conservative 187; Mismatches 401; Indels 108; Gaps 35;

Qy 1 MGASSSTGQRPEPQ-PAGSDTPSELVLSHGCPAAEAS-GAAGPADAD--DPATLPQKNG 56
Db 1 MGASSSTGQRPEPQPGSSSTPAPEPPSGGSAEAPDTPADAIASDPATLPLKNG 60
Qy 57 QLSNVGVAEGQDVHVOENQEQ-----EEEVDEVDYGOREESEDPVREKDR 102
Db 61 QLSITNGVAEODELSLOBGDNGQKALNGQALNSQEEBEVITYEVQQRDSEVSEKRS 120
Qy 103 VEEMANSTAVEDITTKQOETSEIIEOIPASENNVEMVQPAESQANDVGFKKVFRVG 162
Db 121 DKEMATSAVVDITDDQGEENRN-IEQIPSESNLIELNQPTESQANDLIGFKKVFVFG 179
Qy 163 KFTVKKDKNEKSTVOLLTYKQDEGGAELASVAGCHOEPSVTTAVGESKSELSKOS 222
Db 180 KFTVKKDKTEKPTVOLLTYKQDEGGA--AGAGDHODPSTL--GAGGAASKSESEKOS 234
Qy 223 TEKQEGTLKQESSTELPLQESQAAEBAKDEGEKQEKPTSPSPSPNSSETTS 282
Db 235 TEKBEETLKREQSHAEISPPAESQAV-EECKEGBEKEKEPEKSAESPTSPVTSSTGS 293
Qy 283 SFFKFTHGAMGMRKTKTSFKKSKEDLETAERKKEQBAEKVDEBEKTEBPASBE----- 337
Db 294 TFKKFTQGMAGMRKTKTSFRKPKEDVEVASEKKEQEPKVDTEEDQKAVASKTLASE 353
Qy 338 ----QEPNEDTDQRLADRYKVELPLBDYQGLDEASSEKCAPLATVPEKMEANQ-E 392
Db 354 QAHQPEPAPESAHBEPLSAEYKVELPSEBQVSGSQGSEEPALALAEVPEKLEIHOEB 413
Qy 393 VVAEHAHVSTVEKTEBEOGGGEAGGVVEEGTGESLPEPKLAEPQVPEAPAEELMS 452
Db 414 VVAEHAHVSTVEERTBO-----KTEVEETAGSVPAEELVGMADAPQEBEPKELVKL 465
Qy 453 REMCVSGDHTQTLTDLSPBEKTLKPHBGIVSEVEMLSQGERIKVQSPCLKLPSSGLK 512
Db 466 KETCVSGEDPTQGADLSPDEKVLSPPEGVYSEVEMLSQGRMKVQSPCLKLPSTGLK 525
Qy 513 KLSGKKQKGRGGGDEPGEYQHIHIESPESADQKGESSASSPPEPEETTCLEKPLE 572
Db 526 KLSGKKQKGRK-GGGDESGEHTQVPADSPDOEKGESSASSPPEPEETTCLEKGLAE 584
Qy 573 APQGEAEEGTSDGEEKRBEGITPMASFKKVTBPKKVRPSESDEKEELEKVSATLSS 632
Db 585 VQOQGEAEAGTSGEKKRBEVTPMASFKKVTBPKKVRPSESDEKEELEKVSATLSS 644
Qy 633 TDSVSEMODEVKTVGEQKPEEPKRRVDTSVSWEALICVSSKKRARKASSSSDEGGR 692
Db 645 TESTASQEMQEMKGSVEEPKPEEPKRYDTSVSWMALICVSSKKRARRRSSDEBGP 704
Qy 693 TLGDSHRAEASDKREAGTDAVASTOEDQAOGSSSPBEPAGSPSREGGVSTWESFKRL 752
Db 705 AMGCDHQKADAGDKDQETGDIAGSEHPGCGSSSPBEPAGSPTEEGGVSTWESFKRL 764
Qy 753 VTPRKSKSLEEKAEAD-----SVLEQLSTELTEPSEBSWWSIKFIPGRKKRPDQKQO 808
Db 765 VTPRKSKSLEEKSESDSIAGSVESHSTPDTPEPKGEESWISIKFIPGRKKRPDQKQO 824
Qy 809 ATVEDSGVEINEDDPNVPVAVPLSEYNANVEREKME--AQNTLPLQILGAVVSEELS 865
Db 825 APVEDAGFTGANEDSDVPAVAVPLSEYDAVEREKMEAOQAQKGAQPEQKATEVSKELS 884
```



```
Qy 866 KTLVHTVSVAVIDGTRAVTSVEERSPSWISASVTEPLEHTAGBAMPPVVEVTEKDIITA-E 924
Db 885 ESQVHMMAAADAADGTRAAITIEERSPSWISASVTEPLEVEAEAAALTVEEVLEEVIAEE 944
Qy 925 ETPVLQTOTLPGKADHDMVTSEVDFTSSEATATETSEALRTEBTEVTEASGAETTDWVSA 984
Db 945 EPPVTEPLENREARGDTVVSEAELETPAETAETAGPLGSEGTETASAAEETEMVSA 1004
Qy 985 VSQUTSDPTTEATPVOEVESGVLDTEEBEROTQAILQAVADVKBESQVPAT----- 1038
Db 1005 VSQUTSDPTTEATPVOEVESGVLDTEEBEROTQAILQAVADVKBESQVPAT----- 1064
Qy 1039 -QTVORTGSKALEKVEVEEDSEVLASEKEDKVMKPGFVQVQVQVQVQVQVQVQVQV 1097
Db 1065 LQPVQRA-----EAEPEPQAEASGLKEDTVDVVLKVDQAQEAKEPPTQGVQVQVQV 1118
Qy 1098 LE-VPEVTADVH---VATQV-----IKLQOL-MEQAAPVSESSETLTDSETNGSTPLAD 1147
Db 1119 FEKAPQVTESESELVTTQAEFLAGVKQGMWQAIIPDSVETPTDSETDGTSTPAD 1178
Qy 1148 SDTAGTQODETIDSDQSKATAAARVQSVTTEEAATAQKEEPSTLPNNVPAQEEHGEPPG 1207
Db 1179 FDAPGTTQKDEIVEIHEENEVASGTQSGGTEAEAVPAQKERPPA-PSSFVFQEBTKQSK 1237
Qy 1208 -RDVLEPTQOELTAAPVPLAKTEVGEQGEVDMLDGEKV-----EEOEVPVHSG---PNS 1259
Db 1238 MEDTLEHTDKEVSVETVSIILSKTEGTQ--EADQYADEKTKDVPPFEGLEGSIDTGIIVSR 1295
Qy 1260 QKAADVTVYDSEVMGACOBKESTEVOS----LSLEEGEMETDVEKEKETKPEQVSEE 1314
Db 1296 EKVTEVALKGECTEACCKDDALELOSHAKSPSPVERMVVQVREKTEAEETHVNEE 1355
Qy 1315 G-EQETAAPHEGYTKPVLTLDPSSBERGKALGSLGS-PSLPDQDKAGCIEVQVQSLD 1372
Db 1356 KLEHETAVTVSEEVSKQLQTVNPIIDGAKVESLSSGPPCLQGEAEVCTKIQVQSS 1415
Qy 1373 TTVTQTAEAEVKVI--ETVVISSETGESPEVCVGAHLLPAEKSSATGCHWTLOHAEDTVPLG 1430
Db 1416 ASFTLTAAAEKEKVLGETANILETGETLEPAGAHVLEEKSEKNEDEFAAHPGSDAVPTG 1475
Qy 1431 PESQAESPIIIVTPAESTLHPDLQGEISASQERSERSEEDKPDAGPDAGKSTAEIKVL 1490
Db 1476 PDCQAKTPTVIVSATTKGLSSDLEGGKTSLSKWKSEDEVQVACQEV--KVSVAISDL- 1532
Qy 1491 KAEP--ILELESKNKIVLNVIOAVDQFART-ETAPETHAYDSQTVQVACRLDSREP 1547
Db 1533 --EPENGILELETKSKSLVQNIQTAVDQFVRTETETATMLTSELQTAHVIKADSDAG 1590
Qy 1548 RCWTYK-----MKDAMKHVPQPPREDLQVLTVLEA 1577
Db 1591 QETEKEGEEPOASAQDETPTITSAKESESESTAVGQA 1625
```

RESULT 13

```
PCT-US01-08631-51377
; Sequence 51377, Application PC/TUS0108631
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-049
; CURRENT APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 51377
; LENGTH: 1795
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
```

```
; NAME/KEY: DOMAIN
; LOCATION: (415)...(463)
; OTHER INFORMATION: G-protein gamma subunit profile domain identified by eMATRIX,
; OTHER INFORMATION: accession number BL50058, p-value=5.557e-09, raw score of 27.23
; NAME/KEY: misc feature
; LOCATION: (1)-(1795)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
PCT-US01-08631-51377
```

Query Match 52.8%; Score 4262.5; DB 1; Length 1795;

Best Local Similarity 57.5%; Pred. No. 8.2e-237;

Matches 951; Conservative 190; Mismatches 410; Indels 103; Gaps 33;

Qy 1 MGAGSSTEQRSPEQ-PAGSDTPSELVLSGHGPAEAS-GAAGDPADA--DPATKLPQKNG 56

Db 9 MGAGSSTEQRSPEQPGSGSTPAEPSPGGPSAEAPDTTADPAIAASDPATKLLQKNG 68

Qy 57 QLSNVNGVAEQGVHVVQENQEQ-----EEVVDEVDQGESDVEKDR 102

Db 69 QLSSTINGVABQDELSTQEGDLNGKALNGQALNSQEEEVIVRQVDHIDSDVSDRDS 128

Qy 103 VEGMAANSTAVEDITKDGQEBTSEIIQIPASBNVNMVQPAESQANDVGFKKVFKFVG 162

Db 129 DTEWTKSAVVHDIITDDQEEPTDIIQIPSSSENLEELTQPTESQANDIGFKVFKFVG 198

Qy 163 FKFTVKDKNKEKSDTVQLLTVKKDEGEGAPASVAGDGHQEPSPVETAVGESAKSESLKQS 222

Db 189 FKFTVKDKTEKPDVTQLLTVKKDEGEGA---AGAGDHQDPSL--GAGEAASKESEPKQS 243

Qy 223 TEKQEGTLKQEQSSTIPIQAESDQAAEBKADGEGEKQEKTKPSPPSPVNSSETTS 282

Db 244 TEKPEETLKREQSHARISPPAESQAV-EECKEEGEKEQEKPSKSAESTSVTSETGS 302

Qy 283 SFKFFFTGHGAGWKTKTSFKKSKEDDLETAEKKEQAEKVDEEKEKTEPASEE----- 337

Db 303 TFKKFTQGWAGWKTKTSFRKPKEDVEASEKKKEQEKVDTDEEGKARVASEKLTASE 362

Qy 338 ----QEPADTDQARLSADYKVELPLEQVGDLEASSEEKCAPLATEVFDEKMEAHQ-E 392

Db 363 QAHPQPAESAHEPRLSAEYKVELPSEQVSGSQGSEBEPKAPLATEVFDEKIEVHQEE 422

Qy 393 VVAEVHVTVEKTEBEGQGGGAEAGVVGEGTGESIPPEKLABPQVPPQABAEELMKLS 452

Db 423 VVAEVHVTVEERTEEQ-----KTEVEETAGSVPAEELVGMADPQABAEPAKLVKL 474

Qy 453 REMCVSGDHTLTDLSPBEKTLPKHPEGIVSEVMSLQOERIKVOGSLPKLFPSSGLK 512

Db 475 KETCVSGEDPTQGDADUSPDEKVLSPBEGVSVSEVMSLQOERIKVOGSLPKLFTSTGLK 534

Qy 513 KLSGKKQKQKGGGDEEPCGEYQHIHTEPESADEQKGESSASSPEEPETTCLKGPLE 572

Db 535 KLSGKKQKQKGR--CGGDEESGEHTQVPADSPDSQEQKGESSASSPEEPBITCLEKGLAE 593

Qy 573 APDQGAEBEGTSDGKKRKGITPWASFKKMVTPKRVRRRPSDSEBEELEKVKVKSATLSS 632

Db 594 VQODGEAEAGATSDGKKRKGITPWASFKKMVTPKRVRRRPSDSEBEELEKVKVKSATLSS 653

Qy 633 TDSVTSEMDDVKTVEGQKPEEPKERVDTSVSWEALICVGSKKRARKASSSSDDGGPR 692

Db 654 TESTASEMOEEMKGSVEEPKPEPKKVDTSVSWEALICVGSKKRARKASSSSDEGGPK 713

Qy 693 TLGGDSHRAEASAKDEAGTDAVPASTQEQDQAGSSSSPEPAGSPSEGEVSTWSEPKRL 752

Db 714 AMGDDHQKADAEAGKXETGTDGILAGSQEHDPQGGSSSPQAGSPTEGEVSTWSEPKRL 773

Qy 753 VTPRKSKSKLEKAEAD----SSVEQLSTEIPEPSRESVWSIKKFTIPGRKKKADGKQEQ 808

Db 774 VTPRKSKSKLEKSEDSLAGSVEHSTPDPGKESVWSIKKFTIPGRKKKADGKQEQ 833

Qy 809 ATVEDSGPVEINEDDNPVAVPLSEYNAVEREKME---AQGNTELPOLLGAVVSEELS 865

Db 834 APVEDAGPTGANEDDSDVPAVPLSEYDAVEREKMEAAQQAQKSAEQPEQKAATEVSKELS 893

```

QY 866 KTVHTVAVAVIDGTRAVTSVEERSPSMISASVTEPLEHTAGAMPVYEVEKDIIA-E 924
Db 894 ESQVMMMAAAVADGTRAAITIIERSPSMISASVTEPLEVEEAALLTEVLEREVIABE 953
QY 925 ETPVLTQTLPECKADHDMVTSEVDFTSVAVATSEALRTVEVTEASGAETTDMSA 984
Db 954 EPPVTEPLPEKREARGDVTVSSEALTPAVTAALTAGLGAEGGTBAASAEETTEMVA 1013
QY 985 VSQLTDSPTTEEARPVQVESGVLDTEEEERQTAIIQAVADKYKESQVPAT----- 1038
Db 1014 VSQLTDSPTTEEARPVQVEEGGVLDIEQERRTQVLAQVAEKYEBESQLTGTGPEVD 1073
QY 1039 -QTVQRTSKALEKVEVEDESEVLASEKEDVMPKGPQVAGAHLAGSGSTGTQATPS 1097
Db 1074 LQVQRA-----EAEPEEQAEASGLKKEVDVLKVAQAEKTEPFTQKVVGGTTPBS 1127
QY 1098 LE-VEEVTADVDH---VATCCV-----IKLQOL-MEQAVAPESSETLIDSETNGSTPLAD 1147
Db 1128 PEKARQVTESSIESSELVTTQCAFTLAGVSGEMWMOAIPPDSVETPTDSETDSTPVAD 1187
QY 1148 SDTADGTQDETIDSQSKATAAVRQSVTEBEAATAQKEEPTLPNNVPAQEHGEBPG 1207
Db 1188 FAPGTQODEIVEIHEEBEVSAGTQSGTEAAVPAQKERPPA-PSSFVFQEEETKEQSK 1246
QY 1208 -RDVLEPPTQELTAAVPLATVEVQGEVDMIDGEKTV-----EEQEVFVHSG--PMS 1259
Db 1247 MEDTLEHTDEKVESVETVSLSTKTEGTQ--EADQVADKTKVDVFPFEGLEGSIDTGLTVSR 1304
QY 1260 QCAADVTDSEWGVAGCOEKESTEVQS-----LSLEGEMETDYEKEREKTV-DEQVSE 1313
Db 1305 EKVTEVALKGEETBEAECCKDALBELQSHAKSPSPVEEEMVQVYERREYQXQSPTHNVE 1364
QY 1314 EG-EOETAPEHEGTYGKPVLTLDMPSSERKALSGSGS-PSLPDQXACITEVQVQSL 1371
Db 1365 EKLHEHTAVTVEEVSQKLTQTVNVP IIDAKEVSSLESGPPPCQGEBAVCTKIQVQSS 1424
QY 1372 DTTVQTAVAVKVI--ETVVISGEGSEPCGAHLLPKEKSSATGCHTTLQAHETVPL 1429
Db 1425 EASFTLTAAEBEKVIGETANILETGETLEPAGAHVLVEKSKENEDPAHPRGDAVET 1484
QY 1430 GRESQAESIPITVTPAPESTLHPDLOGEISASORERESEEDKPDAGPDADGKESTAIKCV 1489
Db 1485 GPDCAKSTPVVVSATTKKGLSDLEGEKTTSLKMKSDDEVQVACQEV--KVSVAIEED 1542
QY 1490 LKABEILLESKSNKIYLVNIQTAVDQFART-ETAPETHAVDSQTVQVACRLDSREPNR 1548
Db 1543 LEPENGILLETKSSKLVONIQTAVDQFVTRTEETEMLTSELQTAHVIRKADSDAQ 1602
QY 1549 CMTK-----MKDAKMHPPVQPREDLQVLTVLEA 1577
Db 1603 ETEKEGDEPQASADEPTITSAKESSESTAVGQA 1636

RESULT 14
US-60-245-228-307
; Sequence 307, Application US/60245228
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN CYCLASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN CYCLASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000878
; CURRENT APPLICATION NUMBER: US/60/245,228
; CURRENT FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 630
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 307
; LENGTH: 1779
; TYPE: PRT
; ORGANISM: HUMAN
US-60-245-228-307
Query Match 52.7%, Score 4253, DB 27, Length 1779;

```

```

Best Local Similarity 57.6%; Pred. No. 2,9e-236;
Matches 952; Conservative 184; Mismatches 406; Indels 110; Gaps 33;

QY 1 MGAGSSTBORSPEQ-PAGSDTPSELVLSGHGPAAPASGAADPAD- ----DPATLQ 53
Db 9 MGAGSSTBORSPEQPPGSSSTPAPEPSGGQPMAL-----AAPDPAAETIAMPBATTLLQ 65
QY 54 KNGGLSSVNGVABEGDVHVEENQEQ-----EEEVNVEDVQORESEVRE 99
Db 66 KNGGLSTINGVABEDELSTLQGDNLNGQKALNGALNSQEEBEVITVEGQORSEVSD 125
QY 100 KDRVEMAAANSTAVEDITKQGOEETSEIIEQIPASENNVEMVQPAASQANDVQFKTVFK 159
Db 126 SDK--EWAATSAVYHDITDQGEETPELIIGIIPSESENLELTQPTSQANDIGFKVFK 183
QY 160 FVGRKFTYKDKNEKSDTVOLLTVKQDGBGAAASVGAQDHPSTVETAVGESAKSEL 219
Db 184 FVGRKFTYKDKTEKKPTVOLLTVKQDGBGGA--AGAGHODPSL--GAGEAASKSESEP 238
QY 220 KQTEKQEGTLKQOSSSTELPLQAESDQAAEEAKDGEKQEKPTKSPSPSPVNSE 279
Db 239 KQTEKQEPETLKRQSHAEISPPAESQAV--EBCKEGEBEKQEKPSKASPTSPVNSE 297
QY 280 TTSSFKEFTTGAMGMRKKTSTFKSKEDDLETAERKQEAKEKVEBEKEKTEBPASEE-- 337
Db 298 TGSFTKFFFTQGMAGMRKKTSTFRPKEDEVASAKKQEBEKVDTBEDGAEVASEKLT 357
QY 338 -----QEPADTDQARLSADYKVELPLEDQVGLDPAASSEKCAPLATEVPEKMEAH 390
Db 358 ASEQAHPEPAPESAHPEPLSLAEYKVELPSEBQVSGSQGPSEEPAPLATEVPEKIEVH 417
QY 391 Q-EVVAEYHNVTEKTEBEOGGGEABEGVVEGTGSLPPEKLAEOVQOEAPAEEL 449
Db 418 QEBVAEYHNVTEBETEEQ-----KTEVEETAGSVPAEELVENDAEQOEAPAEEL 469
QY 450 MKSREMCVSGDHTQLTDLSPBEKTLPHGPEGIYSEYEMLSQERIKVQSGPLAKLTPSS 509
Db 470 VKLKECTVSGSDPTQAGADLSPDEKVLKSPREGVSEYEMLSQERIMVQSGPLAKLTPST 529
QY 510 GLKILSGKKQKQKGGGGDEEPGEVQHIHTSPSPADQEKESGASSPDEPETTLEKG 569
Db 530 GLKILSGKKQKGR--GGGDEESGHTQVPADSPQOEKQESGASSPDEPETTLEKG 588
QY 570 PLEAPQDEABEGTTSDQEKREGITTPASFKKVTTPKRRRSESPSEBELKVVAT 629
Db 589 LAEVQODEABEGATSDQEKREGTTPASFKKVTTPKRRRSESPSEBELKVVAT 648
QY 630 LSTSTVSEMODQEVKTGGEQKPEEPKRRVDTSVSEALICVSSSKKRAKASSDDEG 689
Db 649 LSTSTVSEMODQEKGVPEPKPEPKRKVDTSVSMALICVSSSKKRAKASSDDEG 708
QY 690 GPTLIGDSHRAEASKDKAAGTAVPASTQEDQAOQSSSPPEPAGSPSEGEVSTWSEF 749
Db 709 GPKAMGSDHQKADKXKQKGTGTGIILAGSQEHDPQSSSPPEQAGSPTEGEGVSTWSEF 768
QY 750 KRLVTPRRKKSXKLEEKED-----SVYQLSTELPSPSEEWVSIKRTIPRRKRRADGX 805
Db 759 KRLVTPRRKKSXKLEEKEDSIIAGSVHSTPDEPGKEEWSVSIKRTIPRRKRRADGX 828
QY 806 QEOATVEDSGVPEIINEDPNVPAVPLSEYNAVEREKME---AQONTPLPOLLGAVVSE 862
Db 829 QEOAVEDAGPAGNEDSDVPVAVPLSEYDAVEREKMEAOQKSAQOPQKATVSK 888
QY 863 ELKTLVHTVAVAVIDGTRAVTSVEERSPSMISASVTEPLEHTAGAMPVYEVEKDI 922
Db 889 ELSEEQVMMMAAAVADGTRAAITIIERSPSMISASVTEPLEVEEAALLTEVLEREVI 948
QY 923 A-EETPVLTQTLPECKADHDMVTSEVDFTSVAVATSEALRTVEVTEASGAETTDMS 981
Db 949 AEEPPVTEPLPENRR--DTVVSEALTPAVTAALTAGLGAEGGTBAASAEETTEM 1006
QY 982 VSAVSQLTDSPTTEEARPVQVESGVLDTEEEERQTAIIQAVADKYKESQVPAT--- 1038

```

```
Db 1007 VSAVSQTLTSDPTTEATPQVEGGVPIEBEQERRTOEVLQVAEKVEESQLPGTGGP 1066
Qy 1039 ----QTVORTGSKALEKVEEVEEDSEVLASEKEKDVMPKGPVQVQAGAEHLAQGSETGOAT 1094
Db 1067 EDVLQPVQRA-----EAERPEEQAEASGLKKTDDVLKVDQAQEAKEPFTQGVVGGTT 1120
Qy 1095 PESLE-VPEVTADVDH---VATCOV-----IKLQOL-MEQAVAPESSETLTDSTNGSTP 1144
Db 1121 PESLEKAPQVTTESLESSELVTTQCAETLAGVKSQEMVMEQAIPPDSEVETPTDSETDGSSTP 1180
Qy 1145 LADSDTADGTQOQDETIDSQDSKATAAARQSVQVTEEEAATAQKEPSTLPNNVPAQEEHGE 1204
Db 1181 VADFDAPCTTKQKDEIVEIHEENEVASGTQSGTAEAAVPAQKERPPA-PSXFFVQEETKE 1239
Qy 1205 EPG-RDVLPTQOELTAAAVPVLAKTEVQGEVDWLDGKVK-----EEOEVFVHSGPN 1258
Db 1240 QSKMEDTLEHTDKEVSVETVSIILSKTEGTQ--EADQVADKTKDVPFPEGLEGSIDTGIT 1297
Qy 1259 SOKAADVTYDSEVMGVAGCOEKESTEVSQ-----LSLEEGEMETDVEKEKETKPEQVSE 1313
Db 1298 VSREKVTVEKGEGTEEAECCKDDALELQSHAKSPSPVEREMVQVVEREKEATEAEPTHVNE 1357
Qy 1314 EG-EQETAAPHEGTYGKPVLTLDMPSSERKALGSLGGS--PSLPDODKAGCIEVQVQSL 1371
Db 1358 EKLEHETAVTVSEEVSKOLLQTVNPIIDCAKEVSSLEGSPPCLGQEEAVCTKIQVSS 1417
Qy 1372 DTTVTQTAAAEVKVETVVISSETGESPECVGAHLLPAEKSSATGGHWTLOHAEDTVPPLG 1431
Db 1418 EASFTLTAABEEKVLGETANITGETLEPAGAHVLVEKSEKNEKEDFAAHGPEDAVPTGP 1477
Qy 1432 ESQAESPIIIVTPAPESTLHPDLOGETISASORERSEEDKPDAGPDADGKESTAIKVLK 1491
Db 1478 DCQAKSTFVIVSATTKKGLSSDLEGEKTTSLKWKSDDEVQVACQEV--KVSVAIEEDL- 1534
Qy 1492 AEPEILELESKNKIVLNVICTAVDQFART-ETAPETHAYDSOTQVPAACRLDSREPNCRW 1550
Db 1535 -EPENGILETKSSKLQVNIITQAVDQFVRTTEETATEMLTSELQTAHVIKADSDAQOET 1593
Qy 1551 TK-----MKDAKMKHPVPQPREDLQVLTVLEA 1577
Db 1594 EKEGEDPQASAOQDETPTSKESESESTAVGQA 1625

RESULT 15
US-09-758-442-562
; Sequence 562, Application US/09758442
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PM031
; CURRENT APPLICATION NUMBER: US/09758,442
; CURRENT FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; NUMBER OF SEQ ID NOS: 710
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 562
; LENGTH: 732
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (67)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (162)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (164)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
```

```
; LOCATION: (464)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-09-758-442-562
```

```
Query Match 9.7%; Score 780; DB 21; Length 732;
```

```
Best Local Similarity 38.6%; Pred. No. 5.8e-36;
```

```
Matches 221; Conservative 87; Mismatches 220; Indels 44; Gaps 17;
```

```
Qy 1039 QTVORTGSKALEKVEEVEEDSEVLASEKEKDVMPKGPVQVQAGAEHLAQGSETGOATPESL 1098
```

```
Db 3 QPVQRA-----EAERPEEQAEASGLKKTDDVLKVDQAQEAKEPFTQGVVGGTTPESEF 56
```

```
Qy 1099 E-VPEVTADV---DHVATCOV-----IKLQOL-MEQAVAPESSETLTDSTNGSTPLADS 1148
```

```
Db 57 EKAPQVTTESIXSSELVTTQCAETLAGVKSQEMVMEQAIPPDSEVETPTDSETDGSSTPADF 116
```

```
Qy 1149 DTADGTQOQDETIDSQDSKATAAARQSVQVTEEEAATAQKEPSTLPNNVPAQEEHGEHPG- 1207
```

```
Db 117 DAPCTTKQKDEIVEIHEENEVASGTQSGTAEAAVPAQKERPPA-PSXFXFOEETKEOSKM 175
```

```
Qy 1208 RDVLEPTQOELTAAAVPVLAKTEVQGEVDWLDGKVK-----EEOEVFVHSG--PNSQ 1260
```

```
Db 176 EDTLEHTDKEVSVETVSIILSKTEGTQ--EADQVADKTKDVPFPEGLEGSIDTGITVSRE 233
```

```
Qy 1261 KAADVTVDSVMGVAGCOEKESTEVSQ-----LSLEEGEMETDVEKEKETKPEQVSEEG 1315
```

```
Db 234 KVTVALKGEGTBEAECKDDALELQSHAKSPSPVEREMVQVVEREKEATEAEPTHVNEEK 293
```

```
Qy 1316 -EOETAAPHEGTYGKPVLTLDMPSSERKALGSLGGS--PSLPDODKAGCIEVQVQSLDT 1373
```

```
Db 294 LEHETAVTVSEEVSKOLLQTVNPIIDCAKEVSSLEGSPPCLGQEEAVCTKIQVQSSSEA 353
```

```
Qy 1374 TVTQTAAAEVKVI--ETVVISSETGESPECVGAHLLPAEKSSATGGHWTLOHAEDTVPPLG 1431
```

```
Db 354 SFTLTAABEEKVLGETANITGETLEPAGAHVLVEKSEKNEKEDFAAHGPEDAVPTGP 413
```

```
Qy 1432 ESQAESPIIIVTPAPESTLHPDLOGETISASORERSEEDKPDAGPDADGKESTAIKVLK 1491
```

```
Db 414 DCQAKSTFVIVSATTKKGLSSDLEGEKTTSLKWKSDDEVQVACQEV--KVSXAIEEDLE 471
```

```
Qy 1492 AEPEILELESKNKIVLNVICTAVDQFART-ETAPETHAYDSOTQVPAACRLDSREPNCRW 1550
```

```
Db 472 PENGILETKSSKLQVNIITQAVDQFVRTTEETATEMLTSELQTAHVIKADSDAQOET 531
```

```
Qy 1551 TK-----MKDAKMKHPVPQPREDLQVLTVLEA 1577
```

```
Db 532 EKEGEDPQASAOQDETPTSKESESESTAVGQA 563
```

Search completed: December 13, 2002, 00:22:46

Job time : 339 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 13, 2002, 00:13:41 ; Search time 85 Seconds
(without alignments)
1253.621 Million cell updates/sec

Title: US-09-902-432-4
Perfect score: 8073
Sequence: 1 MGAGSSTEQRSPQAGSDT.....AWAQRKCLRLQLKAPVSK 1596

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 237906 seqs, 6676526 residues

Total number of hits satisfying chosen parameters: 237906

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents_AA_New.*
1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4335	53.7	1844	5	US-09-724-676-61346
2	4335	53.7	1844	5	US-09-724-676A-61346
3	4294	53.2	1781	1	PCT-US02-04915-44
4	4098	50.8	1685	5	US-09-724-676-61347
5	4098	50.8	1685	5	US-09-724-676A-61347
6	429	5.3	2665	6	US-10-203-138A-11282
7	401.5	5.0	4368	5	US-09-724-676-73643
8	401.5	5.0	4368	5	US-09-724-676A-73643
9	400	5.0	2642	5	US-09-724-676-55851
10	400	5.0	2642	5	US-09-724-676A-55851
11	400	5.0	2722	5	US-09-724-676-55868
12	400	5.0	2722	5	US-09-724-676A-55868
13	400	5.0	2915	5	US-09-724-676-55849
14	400	5.0	2915	5	US-09-724-676A-55849
15	400	5.0	2915	5	US-09-724-676A-55850
16	400	5.0	2915	5	US-09-724-676A-55850
17	400	5.0	2929	5	US-09-724-676-55846
18	400	5.0	2929	5	US-09-724-676A-55846
19	400	5.0	2940	5	US-09-724-676-55847
20	400	5.0	2940	5	US-09-724-676A-55847
21	400	5.0	2940	5	US-09-724-676A-55848
22	400	5.0	2940	5	US-09-724-676A-55848
23	400	5.0	2995	5	US-09-724-676-55866
24	400	5.0	2995	5	US-09-724-676A-55866
25	400	5.0	2995	5	US-09-724-676A-55867
26	400	5.0	2995	5	US-09-724-676A-55867

27	400	5.0	3009	5	US-09-724-676-55863	Sequence 55863, A
28	400	5.0	3009	5	US-09-724-676A-55863	Sequence 55863, A
29	400	5.0	3020	5	US-09-724-676-55864	Sequence 55864, A
30	400	5.0	3020	5	US-09-724-676A-55864	Sequence 55864, A
31	400	5.0	3020	5	US-09-724-676A-55864	Sequence 55865, A
32	400	5.0	3020	5	US-09-724-676A-55865	Sequence 55865, A
33	388	4.8	1325	6	US-10-203-138A-12693	Sequence 12693, A
34	382	4.7	2014	5	US-09-724-676-49425	Sequence 49425, A
35	382	4.7	2014	5	US-09-724-676A-49425	Sequence 49425, A
36	380.5	4.7	2101	5	US-09-724-676-49404	Sequence 49404, A
37	380.5	4.7	2101	5	US-09-724-676A-49404	Sequence 49405, A
38	380.5	4.7	2101	5	US-09-724-676-49419	Sequence 49419, A
39	380.5	4.7	2101	5	US-09-724-676A-49420	Sequence 49420, A
40	380.5	4.7	2101	5	US-09-724-676-49421	Sequence 49421, A
41	380.5	4.7	2101	5	US-09-724-676A-49422	Sequence 49422, A
42	380.5	4.7	2101	5	US-09-724-676-49423	Sequence 49423, A
43	380.5	4.7	2101	5	US-09-724-676A-49424	Sequence 49424, A
44	380.5	4.7	2101	5	US-09-724-676A-49404	Sequence 49404, A
45	380.5	4.7	2101	5	US-09-724-676A-49405	Sequence 49405, A

ALIGNMENTS

RESULT 1
US-09-724-676-61346
; Sequence 61346, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724, 676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 61346
; LENGTH: 1844
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-61346

Query Match 53.7%; Score 4335; DB 5; Length 1844;
Best Local Similarity 58.2%; Pred. No. 2.1e-248;
Matches 962; Conservative 188; Mismatches 401; Indels 102; Gaps 32;

Qy	1	MGAGSSTEQRSPQAGSDT	PSLVLGSHGPAEAS	GAAGDPADA	-DPATKLQKNG	56
Db	62	MGAGSSTEQRSPQAGSDT	PSLVLGSHGPAEAS	GAAGDPADA	-DPATKLQKNG	121
Qy	57	QLSSVNGVAGQDVHVGQ	EEVDEVDVQRESE	DVREKOR	102	
Db	122	QLSTINGVAEQDELSQL	QEGDLNGKQALNG	QSEEEVITVEVQR	SDVSERDS	181
Qy	103	VEEMAANSTAVEDITK	QOSETSIIEQIPASE	NNVEMVQPAESQ	ANDVCFKVPFVG	162
Db	182	DKEMATKSAVVHDITD	QOETPEIIEQIPSE	SNSLBELTQPTES	QANDIGFKVPFVG	241
Qy	163	FKFTVKDKNEKSDTV	OLLTVKKDEGGA	ASVAGAGHQP	SVETAVGESAKSELKQS	222
Db	242	FKFTVKDKNEKSDTV	OLLTVKKDEGGA	ASVAGAGHQP	SVETAVGESAKSELKQS	296
Qy	223	TEKQEGTLKQESSTE	IPLOAESQAAEEBA	KDEGEKEKEPTK	SPSPSPVNSSTTS	282
Db	297	TEKPEETLKQESHAE	ISPPAESQAV	-EECKEESGEKEKE	PSKSAESPTSPVTS	355
Qy	283	SFKKFPFTHGAWGR	KTKTSFKKSKEDDL	ETAKRKEQSAEKV	DEEEKTEPASEE	337
Db	356	TFKKFFTQGWAGWR	KTKTSFKKSKEDDL	ETAKRKEQSAEKV	DEEEKTEPASEE	415
Qy	338	---OEPADTDQARL	ADYKVELPLED	QVDLEASSEKCA	PLATEVDFEKMEHQ	E 392
Db	416	QAHQPEPAESAHP	RLSAEYKVELP	SEBQSGQSPSEK	PAPLATEVDFEKMEHQ	475

```

QY 393 VVAEIVHSTVETKTEBEOGGGAEAGVNVVEGTGSELPPEKLAEPQEVPOEAPEELMK 452
D 476 VVAEIVHSTVETKTEBEO-----KTEVEETAGSVPAEELVENDAEPOEAPEELVKL 527
QY 453 REMCVSGGDHQTLDLSPEEKTLPKHPEGIVSEVEMLSOERIKVQSGPLKLFSSGLK 512
D 528 KETCVSGEDPTQAGADLSDEKVLSPKPGVSEVEMLSOERMKVQSGPLKLFSTGLK 587
QY 513 KLSGKQKQKRGKGDEEPGEVQHITHTESPSADOKESASSPPEPEERTTCLKGP 572
D 588 KLSGKQKQKGR-GGGDEBSGHTQVPADSPQOEOKESASSPPEPEERTTCLKGLAE 646
QY 573 APOGEAEEGTSPGSEKREGITTPWASFVKWTPPKRRRPRESEPKBELKVSATLSS 632
D 647 VQOGEAEEGATSDGEKREBGTVPASTFKWTPPKRRRPRESEKDELVDKVSATLSS 706
QY 633 TDSVTSEMODVKTGEEQKPEEPKRRVDTSVSEWALICVSGSKKRAKASSDDEGGPR 692
D 707 TESTASEMOEEMKGSVEPKPEPRKKTDSVSWALLCVGSSKRRARSGSSDDEGGPK 766
QY 693 TLGGDSHRAEASKDEAGTDAVPASTQEDQAQSSSPPEAGSPSEGEVSTWESFKRL 752
D 767 AMGDHQADEAGKDKETGTGDIILAGSQSSSPQOAGSPTEGEVSTWESFKRL 826
QY 753 VTPRKKSKSKLEKAEAD-----SSVEQLSTEIPRSREBSWSTKRTIPRRKKRKGQ 808
D 827 VTPRKKSKSKLEKSEDSIASGVHSTPDPTEPGKEBSWSTKRTIPRRKKRKGQEQ 886
QY 809 ATVEDSGVEINEDPNYPAYVPLSEYNAVEREKME---AQONTLPOLLGAIVYSELS 865
D 887 APVEDAGTGANEDSDPAPVAVPLSEYNAVEREKMEAOQASAPQOKANTEVSKLS 946
QY 866 KTLVHTVSAVIDGTRAVTSVEERSPSWISASVTEPLEHTAGAMPVVEVTEKDIIA-E 924
D 947 ESQVHMAAAVADGTRATVITIEERSPSWISASVTEPLEHVEALALLTEVLEREVIAEB 1006
QY 925 ETPVLTOTLPREGKADHDMVTSEVDTSEAVTATTSBALRTEETVETASGAETTDMISA 984
D 1007 EPPVTEPLPENREARGTIVVSEALTPREAVTAATAGPLGAEETGESALETTEWISA 1066
QY 985 VSQTLSDPTTEBEATPPOVEVSGVLDTEEEEROTQALLOAVADKKESSQVPAT----- 1038
D 1067 VSQTLSDPTTEBEATPPOVEVSGVLDTEEEEROTQALLOAVADKKESSQVPAT----- 1038
QY 1039 -QTVORTGSKALEKVEVEDESEVLASEKDVMPKGVQEAHLLAQSGSETQCATPES 1097
D 1127 LQPVQRA-----EAEREPEQABASGLKKEPDVVLKVDQAQEAKEPTFPQGVVQGTTPES 1180
QY 1098 LE-VPEVTAADVDH--VATQV-----IKLQOL-MEQAVAPBSSTLLDSENGSTPLAD 1147
D 1181 FEKAPQVETESIESSELVTTQCAETLAGVKSQEMWMOQAIIPDPSVETPTDSETDGTVPAD 1240
QY 1148 SDTADGTQODELTISQODKATPAVROSQVTEEBEATAQKEPSTLPNNVPAQOEBSGEPG 1207
D 1241 FDAEPTTQKDELVEIHENEVASGTQSGTEAEVAPQAKERPRA-PSSFVPOEETKEKESK 1299
QY 1208 -RDVLEPTQOELTAAAVFLAKTEVGOGEVMDLGEKVK-----EEOEVFNHSG-PNS 1259
D 1300 MEDTLEHTDKESVETVSLISTEBGTQ--EADQVADKTKQVFPREGLEGSIDTQITVSR 1357
QY 1260 QKADAVTVDSVMGVAAGQOEKESTEVQ-----LSLEGEKETTIVKEKRETKPQVSEB 1314
D 1358 EKVEEVALKGEETBAECKDQDALQSHAKSPRPVREMVQVQERKEKTEAEPHVAEB 1417
QY 1315 G-EQETAPRHEGTGKVLTLDMSSERKALSGLSG-PSLPRODAGCTEVVOGSLD 1372
D 1418 KLEHTAATVSEVSKOLLQTVNPIIDGAKEVSSLSPSCGLQEAVALCTKIDVQSSSE 1477
QY 1373 TTVTQTAAVEKVI--ETVVISSETGESPEVCAGHLLPAEKSSATGAGHNTLQHAEDTVLG 1430
D 1478 ASFTLTAABEEKVLGTETANILFETGETLEPAGAHVLEKSEKKNEDPRAAHNGEAVTGT 1537
QY 1431 PESQAESIPITVTPAESTLHPDLQGEISAGQRESEEBEDKPADGADGKESIAIEKVL 1490

```

```

D 1538 PDCQAKSTPVIVSATTTKKGJSSDLGEKTTSLKXKSDVEVDQVACQEV--KVSAILEDL 1595
QY 1491 KAEPELLELESKSNKIYLVNIQTAVDOPART-ETAPETHAVDSQTVPAGRUDSREPRNC 1549
D 1596 EPENGILLELETSKLVONITIQFVADQFVTRTEATATMLNSELQTAHVIAKADQDQGE 1655
QY 1550 WTK-----MDAKKHHVPOPREDLQVLYLEA 1577
D 1656 TEKEGEPPQASADDEPTITSAKESESTAVGOA 1688

RESULT 2
US-09-724-676A-61346
; Sequence 61346; Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724, 676A
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 61346
; LENGTH: 1844
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-61346

Query Match 53.7%; Score 4335; DB 5; Length 1844;
Best Local Similarity 58.2%; Pred. No.2.1e-248;
Matches 962; Conservative 186; Mismatches 401; Indels 102; Gaps 32;

QY 1 MGAGSSTGORSPEQ-PAGSDTPSELVLSGHGPAEAS-GAAGPRADA--DPATYLPQKNG 56
D 62 MGAGSSTGORSPEQPREGSSTPAEPSPGGGSPSAEAPDTTADPALAASDPATKLLQKNG 121
QY 57 QLSGVNGVAEGQDVHVOBENOEGQ-----EEBVDEVDVQGBSEDVREKDR 102
D 122 QLSGVNGVAEGQDELSTLQGBDLNGQKALNGQALNSQEBREVTVEVQGBDSEBVSERDS 181
QY 103 VEEMAASTAVEDITTKQGOEETSLIEQIPASNNVVEEMQAPESQANDVGFKKVFPVFG 162
D 182 DKEMATYSAVVDITLDQGEETPELIEQIPSSBSNLELTPQTESQANDVGFKKVFPVFG 241
QY 163 FKFTVKDKNEKSDTVQLLTVKKDGEGAASVAGAGHOEPPSVETAVGASAKSESLKOS 222
D 242 FKFTVKDKNEKSDTVQLLTVKKDGEGA---AGAGHDQPSL--GAGGAASKSESLKOS 296
QY 223 TEKOEGTLKOBQSTETPLQAESQDALEBAKDEBGEKQEKPTKSPSPSSPVNSETTS 282
D 297 TEKEEETLKREQSHAEISPAESQAV-ECKEKEBGEKQEKPEKSPASPTSPVTSSTGS 355
QY 283 SFKKFTTHGAAWGRKKTSPFKSKXKEDLETAEKREKQEAQVDEBKEKTEPASE----- 337
D 356 TEKKFTQGWAGMKKKTSPFKPKEDVEASEKKEQEPKAVDTBEDGALEVASKLTASE 415
QY 338 -----QEPAEPTDQALADYKVELPLEDQVGDLEASSEKCAPLATEVEPEKKAHQ-E 392
D 416 QAHQEPRAESAHNERPLAEYKVELPSEBOVSGSGSEKRPALAEVDEKIEVQOEE 475
QY 393 VVAEIVHSTVETKTEBEOGGGAEAGVNVVEGTGSELPPEKLAEPQEVPOEAPEELMK 452
D 476 VVAEIVHSTVETKTEBEO-----KTEVEETAGSVPAEELVENDAEPOEAPEELVKL 527
QY 453 REMCVSGGDHQTLDLSPEEKTLPKHPEGIVSEVEMLSOERIKVQSGPLKLFSSGLK 512
D 528 KETCVSGEDPTQAGADLSDEKVLSPKPGVSEVEMLSOERMKVQSGPLKLFSTGLK 587
QY 513 KLSGKQKQKRGKGDEEPGEVQHITHTESPSADOKESASSPPEPEERTTCLKGP 572
D 588 KLSGKQKQKGR-GGGDEBSGHTQVPADSPQOEQKESASSPPEPEERTTCLKGLAE 646

```



```
QY 573 APOGEAEEGTGSGEKKBEIGITWASFKKMWTPKKVRRPSESDBKEELEKVKSATLSS 632
Db 647 VQDGEAEEGATSGEKKREGVITWASFKKMWTPKKVRRPSESDBKEDLDKVKSATLSS 706
QY 633 TDSVSEMQDVKVTGEBQPEEKRRVDTSVSWEALICVGSKKRKARKASSSDDEGGPR 692
Db 707 TESTASEMQEMKGSVEEPKPEEKRRKVDTSVSEALICVGSKKRRARRGSSDEEGPK 766
QY 693 TLGDSHRAEASKDKAGTDAVPASTQEQDAQGSSPPEAGSPSGEGVSTWESPKRL 752
Db 767 AMGDHQADEAGDKETGTDGILAGSQEHDPGQGSSSPEQAGSPTEGEGVSTWESPKRL 826
QY 753 VTPKKSLSLEKAEAD----SSVEQLSTELPESRESWSIKKFIPIRRKKRAGDKQEQ 808
Db 827 VTPKKSLSLEKSESDISAGSGVEHSTPDTEPGKESWSIKKFIPIRRKKRPDGQEQ 886
QY 809 ATVEDSGPEINEDDPNPAVPLSEYNAREKME--ACGNTELPQLLCAGVVSVSELS 865
Db 887 APVEDAGTGANEDSDPAVPLSEYDAVEREKMEAAQQAQSAEQPEQKAATEVSELS 946
QY 866 KTLVHTVSAVIDTRAVTSVEERSPSWISASVTEPLEHTAGTAGEAMPPVEEVEKDIIA-E 924
Db 947 ESQVHMAAAVADGTRAATIIERSPSWISASVTEPLEQVEAEAAALLTEEVLEREVIABE 1006
QY 925 ETPVLTOTLPEGKAHDMVTSEVDFTSEAVTATETSEALRTEBVTGASGABETDMVSA 984
Db 1007 EPPVTPELPENREARGDTVVVSEALTEAPEAVTAETAAGPLGAEEGTASAAEETEMVSA 1066
QY 985 VSQLTDSPTDTEETATPQVESGVLDTEEBERQTAILOAVADVAKVEESOVPAI----- 1038
Db 1067 VSQLTDSPTDTEETATPQVEGGVPDIEEQRRITQEVLOAVAEKVKEESQLPTGGPEDV 1126
QY 1039 -QTQRTGSKALEKVVEEEDSEVLASEKEKDVMPKGPVQVQAGAEHLAQGSETQCATPES 1097
Db 1127 LQPVORA-----EAERPEQAEASGLKETTVDVLKDAQEAKEPTFTQGVKVGQTTPES 1180
QY 1098 LE-VPEVTADVDH---VATQV-----IKLQOL-MEQAVAPESSETLTDSETNGSTPLAD 1147
Db 1181 FEKAPQVTEGIESSELVTTTQAEATLAGVKSQEMVMEQAIIPDSVETPTDSTDGSTRVAD 1240
QY 1148 SDTADGTQDDETIQDSKATAAARQSVQTEEBEAATQAKESPSTLPNNVPAQREHGEPEG 1207
Db 1241 FDPCTTQKQDEIIVIHENEVASGTQGGTEAEAVPAQKERPPA-PSFVPQESTKQSK 1299
QY 1208 -RDVLEPTQOEBLTAAPVLAKTEVGOGEVDWLDGKVK-----ERQEVFVHSG--PNS 1259
Db 1300 MEDTLEHTDKEVSVETVSIILSKTEGTQ--EADQYADEKTKDVFPFEGLEGSIDGTIVSR 1357
QY 1260 QKAADVTYDSEVMGACQOEKESTEVQS-----LSLEBEGEMETDVEKEKRTKPEQVSEE 1314
Db 1358 EKVTEVALKGEGTEAECKKDDALELQSHAKSPSPVEREMVQVVEREKTEAEPTHVNEE 1417
QY 1315 G-EQETAAPHEGTYGKPVLTLDMPSSERKALGLSGS-PSLPDQKAGCIEVQVQSLD 1372
Db 1418 KLEHETAVTVEEVSQKLQTVNVPIDGAEVSSLSGSPPCPGQBEAEVCKIQVQSSE 1477
QY 1373 TTVTQTAAVEKVI--ETVVISSETGESPECVGAHLLPAEKSSATCGHWTLOHABDTVPLG 1430
Db 1478 ASFTLTAABAEKVLGETANILETGETLEPAGAHVLEKSSKERNEDFAHPGEDAVPTG 1537
QY 1431 PESQAESPIIIVTAPESTLHPDLOEISASQRESREEDKPDAGPADGKESITAISKVL 1490
Db 1538 PDCQAKSTPVIVSNTATKGLSSDLEGETTSLKWSDEVDRQVACQEV--KVSVAIBEDL 1595
QY 1491 KAEPIELELESKSNKIVLNVLTQADVPART-ETAPETHAVDSQTOVPACRLDSREPNC 1549
Db 1596 EPENGILELETSSKSLVQNIITQAVDQFVRTEETATENMTSELQTAHVIKADSQDAQOE 1655
QY 1550 WTK-----MKDAMKHPVQPPQREDLQVLTVLEA 1577
Db 1656 TEKESGEPOASAQDETPTSKESESESTAVQQA 1698
```

```
RESULT 3
PCT-US02-04915-44
; Sequence 44, Application PC/TUS0204915
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Glynn, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Aziz, Natasha
; APPLICANT: Bos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Angiogenesis, Compositions and
; FILE REFERENCE: 018501-006200PC
; CURRENT APPLICATION NUMBER: PCT/US02/04915
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 09/784,356
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/791,390
; PRIOR FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: US 60/285,475
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/310,025
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/334,244
; PRIOR FILING DATE: 2001-11-29
; NUMBER OF SEQ ID NOS: 230
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 44
; LENGTH: 1781
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-04915-44

Query Match 53.2%; Score 4294; DB 1; Length 1781;
Best Local Similarity 57.9%; Pred. No. 5.4e-246;
Matches 959; Conservative 187; Mismatches 401; Indels 108; Gaps 35;

QY 1 MGAGSSTEQRSPBQ-PAGSDTPSELVLSHGHPAAEAS-GAAGDPADA--DPATKLPQKNG 56
Db 1 MGAGSSTEQRSPBQ-PAGSDTPSELVLSHGHPAAEAS-GAAGDPADA--DPATKLPQKNG 60
QY 57 QLSVNGVAEGDVHVEENQEQ-----EEVVDVEDVQRESESDVREKDR 102
Db 61 QLSTINGVARQDELSLQEGDLNGKQALNGQALNSQEEBEVITVEGQDSESDVSRDS 120
QY 103 VEWMAANSTAVEITKDGQSETSEIIEQIPASNNVMEVQPAESQANDYGFKKVFQVPG 162
Db 121 DKEMATKSAAVVDITDDGQENRN-IEQIPSSSNLEELTQPTESQANDIGFKKVFQVPG 179
QY 163 FKFTVKDKNKEGSDTVOLLTVKKDEGEGBAASVAGAGDHQPSVETAVGESAKSESELKQS 222
Db 180 FKFTVKDKTKEPDTVQLLTVKKDEGEBA--AGAGDHQPSL--GAGEAASKESEPKQS 234
QY 223 TEKQEGTLKQEQSSTEIPLQAESDQAAAEBAKQEGEKQKPTKSPSPSPVNSSTTS 282
Db 235 TEKPEETLKREQSHAETSPPAESQAV-EECKEGBEKEKQEKPSKSAESPTSPVTSETGS 293
QY 283 SFKKFFTHGAGWRKTKTSFKKSKEDDLLETAKEKEQAEKVDEEKEKTEPASEE----- 337
Db 294 TFKKFFTHGAGWRKTKTSFKKSKEDDLLETAKEKEQAEKVDEEKEKTEPASEE----- 353
QY 338 ---QEPAEDTDQARLSADYKVELPLEDQVGLASSEBEKCAPLATEVFEDEKNEAHQ-E 392
Db 354 QAHPQEAESAHEPRLSAEYKVELPSEEQVSGQSPSEKPAFLATEVFEDEKNEAHQ-E 413
QY 393 VVAEVHVTVEKTEEBEGGGEABGVVVEGTGESLPPEKLABEPQEVQPAEPAEELMKS 452
Db 414 VVAEVHVTVEERTEEQ-----KTEVEETAGSVAEELVGMADBPQAEPAELVKL 465
QY 453 REMCVSGDHTQLTDLSPBEKTLPKHPIGIVSEVEMLSQERIKVQGSPLKLFSSGLK 512
Db 466 KETCVSGEDPTQAGDLSPDEKVLSPKPPGVSEVEMLSQERIKVQGSPLKLFSTGLK 525
```

```

QY 513 KLSGKQKQKGGGDEPEYOHITHSPESADOKGSSASPEPEBETTCLKGPLE 572
Db 526 KLSGKQKQKGR-GGGDESGEHTQVPADSPQOEKQSSASSPEPEBETCLEKGLAE 584
QY 573 APQGEAEGETTSQOEKREGITTPWASPKKWTPKKRRRPPESKPEBELEKVSATLSS 632
Db 585 VQOGEAEGETTSQOEKREGITTPWASPKKWTPKKRRRPPESKPEBELEKVSATLSS 644
QY 633 TDSITVSEKQDEKVTGVEEQKPEEPKRRYDTSVSWMALICVSSKKRARKASSDDEGGPR 692
Db 645 TESTSEMGEKMSVVEPKPEPKRYDTSVSWMALICVSSKKRARRRSSDDEGGPR 704
QY 693 TLGDSHRAEASKDXKAGTDAVPASTQODQAGSSSPPEPAGSSESGVSTWMSFKRL 752
Db 705 AMGDHQKADKXGKDEGTGIDILAGSOEHDPQOGSSSPPEAGSPTPEBGVSTWMSFKRL 764
QY 753 VTPRRKSKKLEEKED-----SSVQLSTELERPSREESVSTKFKIPGRKKRQADKQOE 808
Db 765 VTPRRKSKKLEEKEDSISAGSVHSTPDEPGKEESVSTKFKIPGRKKRQADKQOE 824
QY 809 ATVEDSGVEINEDDPNPAVPLSEYNAVEREKME---AQNTLPLQLLGAVYVSELS 865
Db 825 APVEDAGTGANEDSDVPVAVPLSEYDAVEREKMEAQOAGAQPEQKATEVSKELS 884
QY 866 KTLVHTSAVAVIDGTRAVTSVEERSPSWISASVTEPLEHTAGEAMPVVEUTEKDTIA-E 924
Db 885 ESQVHMAAAVADGTRAAITIEERSPSWISASVTEPLEVEAALITEEVLEREVIABE 944
QY 925 EFPVLTOTLPEKKAHDMVNSEVDFTSAAVATESELTRETEVTPKASGAEETDMVSA 984
Db 945 EFPVTEPLPENREARGDVTVSEALTPRAVTAETAGPLGSEEGTVAABETTEMVSA 1004
QY 985 VSQDLSPTTEEARPVQVESGVLDTSEERQTOAILQAVADKVEESQVAT----- 1038
Db 1005 VSQDLSPTTEEARPVQVEGSGVLDIEQERTOEVLQAVAKKESQGLGTGGPEV 1064
QY 1039 -QTVORTSSKALEKVEVEDESEVLASEKQVMPKGPVQEAHGLAOGSTGQATPES 1097
Db 1065 LQPVQRA-----EABREPEQASGLKKEETDVVLKVDQOEAKTEPTQGVKAVGQTPES 1118
QY 1098 LE-VEEVNADVH---VATCOV-----IKLOOL-MEOAABESSELTLDSENGSTPLAD 1147
Db 1119 PEKAQVVESTIESSSELVTTQOETLAGVKSQEMVMEQALPPDSVETPTDSEDTGSPVAD 1178
QY 1148 SPTADQOODETIDSDSKATAAVROSQVTEEEAATAQKEEPTSLPNNVPAOEHEGEBG 1207
Db 1179 PPARGTQKDEIVEIHEHEVNASGTQSGCTEAFAVPAQKERBPA-PSSVFPQEBETKEGSK 1237
QY 1208 -RDVLEPTQOELTAAAVPVLATVEGQSGEVDMLDGEKVK-----EBQEVFVHSG--PNS 1259
Db 1238 MEDTLEHTDEKESVETVSLSTKEGTQ--EADQVADKTKDVPFEGLEGSDTGTIVSR 1295
QY 1260 QKAADVTVDSEVMGAGCOEKESTEVOS-----LSLEBEMETDVEKEKRETKPEOVSE 1314
Db 1296 EKVTVALKGBETSEBECKDDALBELQSHAKSPPSFVEHEMVVQVEREKTEAEPHVNBE 1355
QY 1315 G-EQETAAPEHEGTGKPVLTLDMPSSERKALGSGS--PSLPQDKAGCTIEVOQSLD 1372
Db 1356 KLEHETAVTVSEVSKQLQTVNVPILIDAKAVSSLESGPPCLQGEAFAVCTIKIVQSSSE 1415
QY 1431 PPSQASBPITVTPPAESTLHPDLOGEISASQERSEEDKPDAGDADGKSTAEKTL 1490
Db 1476 PDCQAKSTPVIASATTKGLSSDLBEKKTSLKWSKDEVDQVACQEV--KVSVALEBD- 1532
QY 1491 KAEPE--ITELSKSKNIVLANTVQTAVDQOFAET-ETAPETHAYDSOTVQVACRLDREPN 1547
Db 1533 --EPENGILLETSKSKLVONTIITQAVDQFVTEBETATMLTSELQTAHVAKDSQDAG 1590

```

```

QY 1548 RCMWK-----MKQAKMKHPVQPREDLQVLTVLEA 1577
Db 1591 QETKEGEBEPOASHQODETPTLSAKESSESTAVGQA 1625

RESULT 4
US-09-724-676-61347
; Sequence 61347, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 61347
; LENGTH: 1685
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-61347

Query Match          50.4%; Score 4098; DB 5; Length 1685;
Best Local Similarity 58.4%; Pred. No. 2,1e-234;
Matches 904; Conservative 177; Mismatches 382; Indels 84; Gaps 28;

QY 89 VQGESEEDVREKDVREEMANSTAVEDITYDQGEITSEIILQIPASENNVEMVQPAESQ 148
Db 9 VQGESESEVSESDSDKEMATKSAVAVHIDTQGEETPEIILQIPSSSENLIELTQPTESQ 68
QY 149 ANDYGFKKVFPKVFQPKFTVKKDKNEKSDTVOLLTVKKDEGGEAASVAGADHOBSPVETA 208
Db 69 ANDIGFKVFPKVFQPKFTVKKDKTEKPTVOLLTVKKDEGGA--AGAGDHQPSL--G 123
QY 209 VGEASKESELSKQSTKQEGTLKQOSTTEIPLQAESDQAAEEBAKDEGEKQKEPTKS 268
Db 124 AGEASKESEPKQSTKQEPTEILKQESHAEISPPAESQAN-EECKKEGEKQKEPKS 182
QY 269 PESPSSPVNSFTTSSFKKFTTHGAGWRKKTSPKSKBEDDLETAKKRQEAKEVDEBEK 328
Db 183 AESPTSPVTSSTGTFKFFFTQAGWGRKKTSPKPKXEDVEASEKKEQPEKVDTEED 242
QY 329 EKTEPASEE-----QEPADTDQARLSADYEXIELPLEDVOGLDEASEKQPIA 379
Db 243 GKAEVASEKLTASQOAHPPQEPASAEPRLSAEYKVELPEEVOVSGSQGSEKPPVLA 302
QY 380 TEVDEKEMEAHQ-EVVAEVHVSIVTEKTEEBEQGGGEABGVVAVGTGSLLPPEKLAEPQE 438
Db 303 TEVDEKTEVEVQAEVVAEVHVSIVTEBRTBEQ-----KTEVETAGSVABEIVEMDA 354
QY 439 VPQEAPEAEELMKSRKEMCVSGGHTLTLSPEKTLPKAPBEGIVSEVMSSQBRIVQ 498
Db 355 EPOEAPEAEELVVKETCTVSGEDPTQCADLSPDEKVLJSKPEGVAVSEVMSSQBRMVQ 414
QY 499 GSPKTKLFESSGLKKLSGKQKGRGGGDEPEYOHITHSPESADOKGSSASSPE 558
Db 415 GSPKTKLFTSTGLKLSKQKQKGR-GGGDESGEHTQVPADSPQOEQGBESSASSPE 473
QY 559 EPEETTCLEKPLPAPQDGEAEGTTSQGEKKREGITTPWASFKKWTVPKKVRRPSEBDK 618
Db 474 EPEETTCLEKPLAEVQDGEAEBGATSQGEKKRBCGVTPMAFKKWTVPKKVRRPSEBDK 533
QY 619 EEELEKVSATLSTDSITVSEMDEKVTGVEEQKPEEPKRRYDTSVSWMALICVSSSKR 678
Db 534 EDELDKVKSATLSTESTASSEMGEKMSVVEPKPEPKRYDTSVSWMALICVSSSKR 593
QY 679 ARKASSDDEGGPRTLGDSHRAEASKDXKAGTDAVAVASQODQAGSSSPPEAGPS 738
Db 594 ARGSSDDEGGPRAMGDHQAEDKXGKDEGTGIDILAGSOEHDPGQSSSPPEAGSP 653
QY 739 EGEVSTWESFKRLVTPPKKSKKLEKAED-----SSVQLSTELERPSREESVSTKFKI 794
Db 654 EGEVSTWESFKRLVTPPKKSKKLEKSEDSISAGSVHSTPDEPGKEESVSTKFKI 713

```


QY	1084	LAGOSETGQATPELE--VPEYTAJDVNH---VATQCV-----IKLOOL--MEQVADESSET	1133
Db	1008	FTQCKVAVGQTTPEFEKAPQVOTESIESSELVTTCCQAFILAGKSGQEMWEOAIIPDSDJET	1067
QY	1134	LTDSETNGSTPLADSDTADGTQODETITDSODSKATPAVRQSOVTEBEAATAQKEPSTLP	1193
Db	1068	PTDSELTGTSFTVADFPDAPGTTQDDEIWEIHEENENVASGTQSGTAEAVAPAKERPFA--P	1126
QY	1194	NNVPAOEHGEHPG--RVLEPPTQOELTAAAVPLATATEVGOEGEVDMDGECVVK-----E	1247
Db	1127	SSFPQOETTKQGSMMETTLTHTDKESVEIVSLSTBGTQ--EADQYADETKQVPPFE	1184
QY	1248	BOEYFVHSG--PNSOKADVTTYDSEYGVAGCQEKESTEVQS----ISLEGEKETIYE	1300
Db	1185	GLEGSIPDIGIVSEKTEVALEAKGEGTEBECKDKDADLEIQTAKSPSPYEREMWQVE	1244
QY	1301	KEKETKPEQVSEBG--QOEYTPARBEHTQCKPLTLTDMSSSEBKALGSLGS--SLPQD	1358
Db	1245	REKTEAEETHNHEKLEHETRAVTVSEBVSQKOLQYVNPFIIDAKEVSSLEBSPPCIQD	1304
QY	1359	DKAGCIEYQVOSLDTTVTQTAEAVERKI--ETVVISETGSEPEVCVGAHLTPAKKSATG	1416
Db	1305	EEAVCTKIQVOSSSASFILTPAAAEKCVILGETANIILTGETTELPAGAHVILBEKSSKNE	1364
QY	1417	HWTLOHADPTPLGPBESQASIPITVTPRAESTLHPDLOGELISASQREHSEEDKPDAPR	1476
Db	1365	DFAAHPGSDAVPTPDDCAQSTPVTYSATTKKGJSSDLBEKRTSLMKMSDVIDOVACQ	1422
QY	1477	DADGKESTALEIKVLAKEPELIELESKSNKIVLNVIGTAVDOFART--ETAPETHAYDSQTQ	1535
Db	1425	EV--KVSVAIEEDLEPENGILLETETKSKIKVQNIIGTAVDOFRTETETATEMLTSELQTO	1482
QY	1536	VPACRLDREPRNCMTK-----MKDAMKHPVQPREDDLOVLTVLEA	1577
Db	1483	AHVIAKDSQDAGQETEKEGEPPQASQADEFTPTISAKESBSSTAVGQA	1529

```

RESULT 6
US-10-203-138A-11282
; Sequence 11282, Application US/10203138A
; GENERAL INFORMATION:
; APPLICANT: Molecular Dynamics, Inc.
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: ANALYSIS OF GENE EXPRESSION IN BT 474
; FILE REFERENCE: PB 0004 WO 8
; CURRENT APPLICATION NUMBER: US/10/203,138A
; CURRENT FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 04 February 2000 (04.02.00)
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 26 May 2000 (26.05.00)
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 03 August 2000 (03.08.00)
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 03 October 2000 (03.10.00)
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 27 September 2000 (27.09.00)
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 21 September 2000 (21.09.00)
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 30 June 2000 (30.06.00)
; NUMBER OF SEQ ID NOS: 15438
; SOFTWARE: Molecular Dynamics Sequence Listing Engine
; SEQ ID NO 11282
; LENGTH: 2665
; TYPE: prt
; ORGANISM: Homo sapiens
; FEATURE:

```

```

; OTHER INFORMATION: MAP TO AL034555.2
FEATURE:
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 12
;
; FEATURE:
; OTHER INFORMATION: EST_HUMAN HIT: AU117052.1, EVALUE 0.00e+00
;
; FEATURE:
; OTHER INFORMATION: SWISSPROT HIT: P08640, EVALUE 3.00e-10
US-10-203-138A-11282

Query Match          5 3%: Score 429; DB 6; Length 2665;
Beet Local Similarity 19.6%; Pred. No. 1.3e-16;
Matches 572; Conservative 270; Mismatches 755; Indels 504; Gaps 83;

```

QY	9	ORSEOPAGSP-----TPSELVLSHGHPALPAAAGS-----AAGPAPADABAT-	49
Db	301	QVOSTPEPAKBDLKLSEBVMKMKPKMEKGLSSHVEVVEKGRLLKARKHLKPEBPADGSAVD	360
QY	50	--KL.PQKNGQLSSVNGVAEGQDVHVOENQOGEVEEVEDVQORESEDVREKORVEMA	107
Db	361	LEKLEAKRRFPADSNLKLAEKQKPEVKSSPEMEDARVL-----SKQPDVSSREVI----	411
QY	108	ANSTAVDITPKDGOEENSELIEDIPASENNVEEMVOPAESQANDVGFKKVYKFGFKFTV	167
Db	412	-----LHREGAEKKPYRKELKRES-----KKI-----	435
QY	168	KKDKNKSDTVOLLTVYKQEGEGAEASVAGDHOEPPSVETAVAGESASKESELKOSTEKOE	227
Db	436	-----KLDRLNTVASPKDCQELASISVSGSRPSSDLOARIGELAG-----ESVENQZ	483
QY	228	GTLKQBOGSTEIP-----LOAESQOAAEE--AKDESEKQOEKEPTKSPESPSSVNS	278
Db	484	-----VOSKKPIPEKPOLQOLVLDQOGERDVRKNYCSLDETPERKSSQOEKSHSVNT	538
QY	279	ET-----TSSEKKEFTFMAGMKKTSFPKSK-----BDDLETAER-----	315
Db	539	EKKIGIDIDHQSYRQKQMEGRKQOMEWELAKSEKFPSPKXDVEYRRSLVHVEGKPP	598
QY	316	-----KEGEAEKYD-----EEKEKTEPASEOEPAEDTDOARLSADYEKVEL	358
Db	599	QDVTDSPSPSKKKMMDHVDPDICTKRENRYSRSROISEBSEPTGSGPSVRGSPHED---	655
QY	359	PLEDVGVLBAAS-----EEKCAP-----LATVEDEKMEAHQEVLEVAVHST	401
Db	656	--EDPIGSPRLLSXKSGPKYKDEKVLPSNITVARESLKPNPDDSRREKQMDMAKITLVS	713
QY	402	VEKTEEEOGGGGEAE--GGVVVEGTGESLPPKEL-----AEBOEVPQEAEPAE--	448
Db	714	INSEDELNRWDSQKQADAGFDVSPNSIICKDSLKASVRDLEBGEVPSDSDEGEHS	773
QY	449	-----LMKREMCVGGGHTQULTDLSPE-----EKLTKHPEGIYSE	485
Db	774	HSPPASALYESSRSLFLLRREDEKLEBERDRSLSSJERNKFSFALDXITTPDTKALER	833
QY	486	VEMLSS-----OBRIKVQSGP-----LKLKFPSSGKLKLSG	516
Db	834	AKSLSSSRHEENWSPFLMDSDRFANFRANNOKEKVDASAPRIPSWMKKKKXITDSEGMKD	893
QY	517	KKQKKRGKGDDDEPPEGYQHITHESPEBSADQKGBSSASSPEEPETTCKLEKPLAEAD	576
Db	894	KKEDHK-----EEBOERQELFASRFLHSSIFEDBSKRLQHLRKKEDSDFTSGRYKOT	948
QY	577	GAEFGTTSDEKERRREGITPMAS-FKKMV-----TPKRVRRSESDKEEELKVKYASLS	631
Db	949	BEGANSTI---DSIQEVPVLFHSRFEHLTRMQQKEXKQKKEVQKQKEDTENHPKTPES	1005
QY	632	STDSTVSEMQ-----DEYKTVGEOQBEPBPKRVDTSVSEALICYGSSKGRARAKSS	684
Db	1006	APENKDELKTPPVGPPSVTVVTLBSAPALEKTKTGDKTVEAPLV---TEBKTVTEPATV	1062
QY	685	SDD-----EGGPRLTGGDSHRAE-----ASKDKAGRDVAPASTODODAOGSSPEP	733
Db	1063	ISEAKPASEPAPAVP---BQLEVDVLPQADBPDKAA--MMPAGYER---GSSGQOP	1111

QY 734 ---AGSPGEVSTWE-----SPKRLVTPRKSKSKLEKAE--DSSVE----- 773
Db 1112 PYLAKPTPOASQASQNSVDPDSTQPLSKPAQKSEANEKPKADPAEADAN 1171
QY 774 ---QLSEIEPSRESWVSIKFTIPGRKKRQADQKQEQATVEDSGPVEINEDDPNPAVV 830
Db 1172 QKAEAPSPQPSASED-LEVDPPVAADKPKNKRKSKTPTVQ-AAAVSIVE-----KPV 1224
QY 831 PLSEYNAVERKMAQAGNT--ELPQLGAVVVSBELSKT-----LVHTVSVAVIDGT 880
Db 1225 RKSE--RIDREKLKRSNPRGEAQKLLKMEAEKIRTRASKNSAADLHPEPSLPLSRT 1282
QY 881 RA--VTSV-----ERSPSWISASVTEPLEHTAGEAMPVEEVTEKDIIE--ETPVL 929
Db 1283 RRRNRSVYATMGHNRSP-----VKPEVQ-----PRVTRKRLERLEQEA 1327
QY 930 TOT-----LPEGKADHDMVTSEVDFTSSEAVTATETSEALRTEVTEASGAE-----ETTD 980
Db 1328 PTTPRGRPKTRRADEEENEKEPAETLKPPGWRSPRSOKTAAGGGPOGKKGNEP 1387
QY 981 MVSAVSOLTSPTTEBATPVQEVESGVLDTTEBEROTQAILQAVADKVEESQVPAQT 1040
Db 1388 KVDAT-----RPEATTEVGPOIGKESMEPKAAEEBAGSQKDRDKDAGTDKNPPETAP 1442
QY 1041 VORTGSKALEKVEVEEDS-----EVLASEKED--VMPKGPVQAGAEHL 1084
Db 1443 VE-----VWEKKPAPEKNKSKRGRSRLAVDKSASLKNVDAVSPRGAQAQAGER-- 1495
QY 1085 AQSETG--QATPESLEVPETADVHVATCQVIKLOQLMEQAVAPESSETLTDTSETNGS 1142
Db 1496 -----ESGVAVSPKSPQKEDGLS-----SQLKSDPDVDPKEPEKEDVSASGP 1541
QY 1143 TPLADSDTADGTQOEDTDS-----QDSKATAAVRQS-----QVTEBEAAT 1183
Db 1542 SPEA--TOLAKMELEQAVEHIAKLAESASAAKYADAPGLAPEDRDKPAHQASETELA 1600
QY 1184 AQK---REPSTLPNNVQAQEBH-CEE-----PGRDVLPTQOEL-TAAAVPVLAKEV 1231
Db 1601 AIGSIINDISEPENFAPPYPGESQTDLPQAPAGALOPSEGMETDEAVSGILETEA 1660
QY 1232 QGEDEVLDGCKEVEQEVFVHSGPNSSQKAAADVTYDSEVNGVAGCOEKESTEVQSLSLE 1291
Db 1661 ATESRRPPVNA PDPS-----AGPTDKRAGNSSETSHSVPEAKSGKEVEVTLVRKD 1712
QY 1292 EGETVDEKRETKPEQVSEGEQETAAPHEGTYGKPVLTLDMPSSBGRKALGSLGG 1351
Db 1713 KGRQKTTSRKRNTNKKWA--PVESHVPESNQAQ-----ESPAANEGLTV---Q 1759
QY 1352 SPSPDQKAGCIEVQVSLDITVTQTA-----EAVEKVIETVVISSETGESPECVGAH 1404
Db 1760 HPEAPQEEK-----QSEKPHSTPPQSCDLSKIPSTENSSQBSISVEERTPTKASVPPD 1813
QY 1405 L-----LPAEKSSATGHTWLTQHAEDTVPLGPESQAESIPITVTPA----- 1445
Db 1814 LPPPPQAPVDEEQARFRVHSIIESDPVTPPSD-PSIPITPLDSVTAALKSPPVASGGI 1872
QY 1446 PESTLHPDLOQIESASQRESSEEDKPDAGDADGKE-----STAIEKVLKAEPIELLES 1501
Db 1873 PHQSPPTKVTEWITROEPPRAQSTPSPALPDPTKASDVDTSSSTLRKIL-MDPKYVBSATS 1931
QY 1502 KSNKVLNVLTQAVDQFARTETABETHAYDSQTVQACRLDSREPNCRTWKMDAKMKHP 1561
Db 1932 VTSVTVTAIEPV-----SAAPCLH-----EAPPPVDSKKP-----LEEKTA 1972
QY 1562 VPQPREDLVLTLEANAQPKCLPRQLK-----APVS 1595
Db 1973 VTN-NSEIQASEVLVA-ADKEKVPVIAPIAKTITSVISRMVPS 2011

RESULT 7

US-09-724-676-73643

; Sequence 73643, Application US/09724676

; GENERAL INFORMATION:

; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 73643
; LENGTH: 4368
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-724-676-73643

Query Match 5.0%; Score 401.5; DB 5; Length 4368;
Best Local Similarity 19.9%; Pred. No. 1.2e-14;
Matches 317; Conservative 272; Mismatches 662; Indels 339; Gaps 63;

QY 8 EORSPPQAGSD---TPSELVLSHGHPAAEASGAAGDADPADATPKLPK-----NGQ 57
Db 1374 EERLAEQQAEREERLAEVEAALEKORLAEHAQAQAARE-AKELQORMQEEVVRR 1432
QY 58 LSSVNGVAEQGDVHVOENQEGQEEVVDDVQRESE-----DVREKDRVVEEMAANS 110
Db 1433 EAAVD--AQCKRSIQEELQQLROSSEAEIQAKQAQAERSLRIEIEIRVRLQLEA 1490
QY 111 TAVEDITKQOEETSEI-IEQIPASNNVEMVQPAESQANDVGFKKVFKVGFKFTVKK 169
Db 1491 TERQGGAGELQALRARABEAQAQQAEEARLRQVQDESQKQAQAEVELASRVKA 1550
QY 170 D---KNEKSDTVOLLTKKDEGEGAE-----ASVAGADHQPESVET-----AVGESAKES 217
Db 1551 ETEAAREKQALQALBELQQAEEARLRQAEVBARQVVALETAQRAAEAELOSRA 1610
QY 218 ELKQSTEKQSGTLLKQOQ--SSTEIPLQAE---SPOAAEEBAKDEGEKQEKETKSPSPS 273
Db 1611 SFPEKTAQLERSQEBHVAQUREEARRAQQAQAERAEARELELRWQKANEALR 1670
QY 274 SPVNSETTSFKFFPTHGWAGWRKTSFKKSKEDDLETAEKKEQAEKVDDEBEKTEP 333
Db 1671 LRLQAEVAQOKSL-----AQAAEAKQKEEAERARRRKGAEQA VRORELAQE- 1720
QY 334 ASEQEPAEEDTQARLSADYKVELPLEDQVGLAS--SEKCAPL-----ATEVFDE- 385
Db 1721 LEQRQLAEGTAQORLAAEQELIRLRAETEQQEQORQLLEELARLQREAAAAATQRQEL 1780
QY 386 -----KWEAHQEVV-----AEVHSTVTEKEEGGCGEAGEGVVVEGTGE-----SL 428
Db 1781 EALAKVRAEMEVLLASKARAEESRSTSEKQRL-----EAGAPRELAEEAARLAL 1836
QY 429 PPE-----KLAEPQEVQPAEPAEELMKREMVCVSGGDHTQLTDLSPKPEGIV 483
Db 1837 ABEAKRQLAEBAARQRAE-AERVLAEKLAAI--GEATRL-----KT-----EAEI 1881
QY 484 SEVEMLSQBERIKVQSGPLKLPSSSGLKLSKQKQKRGGGDEEPGEYQHHTESPE 543
Db 1882 ALKEKAENER-----LRLAEDEAFORRLREQAQAHKADIEERLAQ---LRKASDS 1931
QY 544 SADEQKQ--ESSASSPEEPEITCLKGPLEAPQDGEAE-----EGTSPGKKEG 593
Db 1932 ELERQKGLVEDTURQRQVEEELALUKAFKAAQAQAELELGRIRSNAEATLRSKEQ 1991
QY 594 ITWASFKKMVTPKKVRPPSPESDK-----EEELKVKSATLSSTSDTSVSEMQDEVKT 646
Db 1992 AELEAARQALAEERERRREAEERVKSLAAEEAARQKAALEVERLKANV-EEARR 2050
QY 647 VGEQKPEPK-----RRVDTSVSWEALICVSGSKRKARAKSSSDEGCPRTLGG 696
Db 2051 LRERAQESARQLQAQEAQAQKLEAKAHAFV--QQKEQLQOTLQEQSVLDQLRG 2108
QY 697 DSHRAEASKDE-----AGTDAVPASTQ-EQDQAQSSSPPEAGSPSEGVSTWESFK 750
Db 2109 EAEARRAAEEAEARVQAREAAQARRQVEAEARLUKQSAEQQAQAQAAAA----- 2162

```

QY 751 RLVTFRKSKSLKEKADSSVEQLSTEIPRESWSVSIKKFIPIRRKKRQADQKQEQAT 810
Db 2163 -----EKLRKEAEQEAARBAQEAQALRQKQADAMEKHKKFAEQTLRQKQVQELTT 2217
QY 811 VEDSGPVEINEDDPNVPVPLSEYNARE-----KME-----AQONTLPOLLGAVYVS- 861
Db 2218 LR-----LQLEETD-----HQNLLDEBELQRLKAEATEAARQSRQVSEBELFSVRQOM 2264
QY 862 EELSKTLVHTVSVAVIDGTRAVTSVEERSPSWISASVTEPLEHTAGAMPV-EEVTEKD 920
Db 2265 EELSK-----LKAIRIENRALLIRDKONTORFLQEBAEKMKQVAAEBAARLS 2311
QY 921 IIAETPVLTQ-----TLBEGKAHDWTSSEVDFTESEATATETSEALRTEVTEBASGA 976
Db 2312 VAAQGAARLRQLAEBDLAQORLAEMKLEKQVAVQEAIRLKAEBELLQOQKEILAQOEAR 2371
QY 977 EETDWSAVSOLTDSPDPTTEEAATPVOEVESGVLDTEEBERQATLQAVADKVK----- 1030
Db 2372 RLQF-----DEQMAQOLAETQGFQRTLEAERQRLMSAEBRLKLRVAM 2419
QY 1031 EESQVPATQVQRTGSKALEKVEEVEDESVLASEKEKQVMPKQVQEAHEHLAQGSET 1090
Db 2420 SAAQARAEBDAQRFKQAEIGEKHLRTE--LATQEKVTLVQTLRIQROSDHDAE---- 2473
QY 1091 GQATEESLEVEPVTDVH--VATQVITLQOL-MEQAVAPESSETLTDSENGSTPLAD 1147
Db 2474 -----RLREAIAELEKEKEKLOOEALILQKSEMQVQOELQLOETQALQOSFLSE 2525
QY 1148 SPTADGTODETIDSDSKATAVAVQSVTEBEAATAQKEBPSTLPNNVPAEEHGEPEG 1207
Db 2526 KQSL--LQERITIQEKALQ-----EQLFQDEBAKQO-----LAEF--QOQO 2564
QY 1208 RDVLEPTQOELTAAAVPVLAKTEVQGEVQVMDLDEKVKQEOEVPVHSGPNSQKADVTY 1267
Db 2565 QOQMEGERQRLVASMBAERROHBAEG-----VARKOE----- 2598
QY 1268 DSEWNGVACQCEKESTEVSLSIEBEMETDVKEKKEKTRPE-QVSEBEGQETAAPEHES 1326
Db 2599 -----ELQOLEQORQOEBELLAEENQRLREQLQLLEEQHRAALAHSEEV 2642
QY 1327 TYGKPLVLTDMPS-----SERKALGSLGSPSLPDQDKAGCTEV-QVQSL---- 1371
Db 2643 TASQVAATITLNGRDALDGPAAEAPEHSPDGLRKVSAQQLQEGILSABELRLQAGS 2702
QY 1372 DTTVTQTAEAVE-----KVLETVVISETGE-----SPECVGAHLIPA 1410
Db 2703 HTTVDELARREVDVRYLQGRSSIAGLLILKATNEKLSVVALORQLLSP--GTALLILEA 2759
QY 1411 SSATG-----GHWTLOHAEDTVPLGE 1432
Db 2760 QAASGFLDPVRNRRLTVNEAVKEGVGPE 2789

RESULT 8
US-09-724-676A-73643
; Sequence 73643, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 73643
; LENGTH: 4368
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-73643

Query Match 5 0%; Score 401.5; DB 5; Length 4368;
Best Local Similarity 19.9%; Pred. No. 1.2e-14;
Matches 317; Conservative 272; Mismatches 662; Indels 339; Gaps 63;

```

```

QY 8 EQRSPQPADS-----TPSELVLSGHGPAABASGAGDPADAPATKLPOK-----NGQ 57
Db 1374 EERLAEQRAEERERLAEVEALEKQRLABAHQAQAQARE-AKELQOQMOQEEVVRRE 1432
QY 58 LSSVNGVAEGDVHVOENEGQEEVEVDEVDGQRESE-----DYREKDRVEAMANS 110
Db 1433 EAAVD--AQQKRSIQBELQOLROSSSEAFQAKARQAEABERSRLRIEBELRVRLQLEA 1490
QY 111 TAVEDITKDGQETSET-IEQIPASENNVEBMOVPAESQANDVGFKKVFFKVFGEKFTLVK 169
Db 1491 TERQGAEGEHLQALRAAEAEAEQKQQAQGEAEERLRRQVODESORQQAEEVLASRVKA 1550
QY 170 D--KNEKSDTVQVLTLYKKQEGEAE-----ASVAGADHGPESVET-----AVGSASKE 217
Db 1551 ETEAAREKORALQALBELRLQAEERLRLOAEVERARQVVALETQORSHEALQSRDA 1610
QY 218 ELKOSTEKEGTLKQOQ-SSTEIPLQAE--SDQAEEKADEBGEKQEKPTSPSPS 273
Db 1611 SFAEKTQRLERSLOEBHVAVQAQLEAEARRAQQAABEABARELEERLQKLANALR 1670
QY 274 SPVNSSETTSFKKPFTHGMAKWRKTSFKSKEDDLETAERKKEQEAQVDEEKEKTEP 333
Db 1671 LRLQAEVVAQOKSL-----AQAEAEKQEBAREARRRGAEBQAVQRELAQOE- 1720
QY 334 ASEQEPREDTDQRLSADYKELVPLEDQVGDLEAS-SEKCAPL-----ATEVFDE- 385
Db 1721 LEKORQLAEGTAQORLAEOQLIRLRAETEGEQEOQRLLEBELRLRQEAANAATQKQEL 1780
QY 386 -----KMEAHQEVV-----AEVHSTVEKTEEBEGGGGAEAGVVEGTGE-----SL 428
Db 1781 EAEALKVAHEEVLASAKAEESRSTSEKQRL-----EAEGRPRELAEEAARLAL 1836
QY 429 PPE-----KLAEPQEPQEAAPAEELMKSRMVCVSGDHTQTLSPBEKTLPHNPEGIV 483
Db 1837 ABEAKRQRLAEEDAAARQAE-AERVLAEKLAAL--GEATRL-----KT-----EAEI 1881
QY 484 SEVENLSQERIKYQSGSLKTLFSSSGIKLSGKKQKQKGGGDEBERGEVQHIHTSEPE 543
Db 1882 ALKEKAENR-----LRLAEDAEAFORRLLEQAAQKHADIEERLAQ--LRKADS 1931
QY 544 SADEQKQ--ESSASSPEEPETTCLEKPLPAQDGEAE-----EGTTSDEKKREG 593
Db 1932 ELERKGLVEDTLQRQRYEEELIATKASPEKAAAGAELELEIGRIRSMNEDTLRSQEQ 1991
QY 594 ITWASFKKQVTPKKRVRRPESDK-----EELEKVKATISSTDSTVSEWODEVKT 646
Db 1992 AELEAARQRLAEERERRRREAEERVQKSLAEAEBAARQKALAEVRLKANY-EEARR 2050
QY 647 VGEQKPEEPK-----RVDTSVSWEALLICVSSKKRAKASSDDEGGPRTLG 696
Db 2051 LREAREQSSARQOLAQEAQRLQAEKHAFAV--QOKEQLQOTLQOQOSVIDQRLG 2108
QY 697 DSHRAEASDKQE-----AGTDVAPASTQ-EQDAQSGSSPEPASPESGEGVGTWESFX 750
Db 2109 EAEARRARAEBEARKVAEEBAQAQRQVEAEERIKSABEQAAQAAQAAAA----- 2162
QY 751 RLVTFRKSKSLKEKADSSVEQLSTEIPRESWSVSIKKFIPIRRKKRQADQKQEQAT 810
Db 2163 -----EKLRKEAEQEAARBAQEAQALRQKQADAMEKHKKFAEQTLRQKQVQELTT 2217
QY 811 VEDSGPVEINEDDPNVPVPLSEYNARE-----KME-----AQONTLPOLLGAVYVS- 861
Db 2218 LR-----LQLEETD-----HQNLLDEBELQRLKAEATEAARQSRQVSEBELFSVRQOM 2264
QY 862 EELSKTLVHTVSVAVIDGTRAVTSVEERSPSWISASVTEPLEHTAGAMPV-EEVTEKD 920
Db 2265 EELSK-----LKAIRIENRALLIRDKONTORFLQEBAEKMKQVAAEBAARLS 2311
QY 921 IIAETPVLTQ-----TLBEGKAHDWTSSEVDFTESEATATETSEALRTEVTEBASGA 976
Db 2312 VAAQGAARLRQLAEBDLAQORLAEMKLEKQVAVQEAIRLKAEBELLQOQKEILAQOEAR 2371

```



```
QY 977 ETTDMVSAVSQLTSDPTTEATPVQESGVLDTTEEROTQALQAVADKVK----- 1030
Db 2372 RLQE-----DKEQMAQALAEQTQGFORTLAEORQRLQEMGAERLKLRAVEM 2419
QY 1031 EESQVPAQTQVGTGSKALEKVEEVEDSEVLASEKDKVMPKGPVQAGAEHLAQGET 1090
Db 2420 SRAQARAEADAQPRKQAEIEGKLRTE--LATQKVTLVQTLEIQROQSDHAE---- 2473
QY 1091 GOATPESLEVPVTDVHD--VATCQVIKLOOL-MEQAVAPESSETLTDSTNGSTPLAD 1147
Db 2474 -----RLREAIAELEREKEKLQAEKLLQLKSEEMQTVQOBLLQETOALQOCSFSE 2525
QY 1148 SDTADGTQOEDTIDSDQSKATAAARQSVQTEEEAATAQKEEPSLTPNNVPAQBEHGE 1207
Db 2526 KDSL--LQRRFPISQEKAL-----BOLFODEVAKAQ-----LREE--QORQ 2564
QY 1208 RDVLEPTQOELTAAPVPLAKTEVQGEVDWLDGKGVKQBEQEVFVHSGPNSQKAADVTY 1267
Db 2565 QQQMEQERQRLVASMEEARROHEAEG-----VRRKQE----- 2598
QY 1268 DSEVMGAGQCKESTEVQSLSLEEGEMETDVEKEKRETPKPE-QVSBEGQETAAPEHEG 1326
Db 2599 -----BLQLEQOQRRQOQELLABENQRLREQLLLEEQHRAALAHSEEV 2642
QY 1327 TYGKPVLTLDMP-----SERKALGSLGSGSPSLPDQDKAGCIEV-QVQSL--- 1371
Db 2643 TASVAATKTLPNGRDALDGPAAEPEHSPDGLRRKVSQARQLQEGITLSAEELQRLAQG 2702
QY 1372 DTTVTQTAEEV-----KVIVTIVISETGE-----SPECVGAHLPAEK 1410
Db 2703 HTTVDELARREDVPHYLGQRSSIAIGLLLKATNEKLSVYALQORLLSP---GTALILLEA 2759
QY 1411 SSATG-----CHWTLOHAEDTVPLGPE 1432
Db 2760 QAASGFLDLPVNRRLTVNEAVKEGVGPE 2789
```

RESULT 9

```
US-09-724-676-55851
; Sequence 55851, Application US/09724676
; GENERAL INFORMATION:
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724, 676
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 55851
; LENGTH: 2642
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-55851
```

```
Query Match 5.0%; Score 400; DB 5; Length 2642;
Best Local Similarity 19.1%; Pred. No. 6.8e-15;
Matches 306; Conservative 304; Mismatches 663; Indels 330; Gaps 58;
```

```
QY 43 ADADPATKLPQKNGQSLSVNGVABQGVHVOEENQEQEVEEVDVDEGVQESBDVREKDR 102
Db 738 AEARQRRLDYES-QTAHDNLLTEQ--IHSLSIEAKSKDKVI---EVLQNELDDVQ---- 787
QY 103 VEEMAANSTAVEDITKDGQETSIL---EQIPASENNVEEMVQPARSQANDVGFKKVKF 159
Db 788 -LQFSEQSTLIRLSQSLQNKSESVLEGAERVRIHSSKVEELSQAQKELEI----- 839
QY 160 FVGEKFTVKDKNEKSTVQLLVTKDEGEAGASVAGDHQEPSVTAVGESAKSESEL 219
Db 840 -----TKMD-----QLLEKRDVETLQOTIEBKDQQVTEISFSMTKVMQVQLNEE 884
QY 220 KOSTKEQEGTLKQFSSTETPLQAESDQAAEEFAKDEGEKQEKPTKSPESPSPVNSE 279
Db 885 KFSLGVEIKTLKE-----QLNLSRABEAKKEQVEEDNEVSSGLKQNYDEMSPAQISKE 939
```

```
QY 280 TTSSFYKFPFTHGAWGRKRTTSFKKSK-----EDLETAERKRKQEQARKVDREKEKT 331
Db 940 E-----LOHEFDLLKKENEQRKRLQALINRKKELLQVRSLBELANLKDSEKIEP 992
QY 332 EPASEQEPAPEDTDQARLS-----ADYEKVELPLEQVQGLEA-----SSBEKCAPLA 379
Db 993 LSTERGEVEEDKENKEYSEKCVTSKQCEIETLYKOTISEKEVELQHIRKDLSEK---LA 1049
QY 380 TEVFDKMEAHQSVAAEVHVSHTVEKTEEOGGGGEAGGVVVVEGTGESLPPEKLAEPQEV 439
Db 1050 AE-----EQFQALVKQMNQTLQDKTNQ-----IDLLAQAEISENQAI 1085
QY 440 PQBAEPAEELMKREMCVSGDHTQLTD--LSPBEKTLPHKPIGIVSEVEMLSQERI 495
Db 1086 IQK-----LITSNTDASDGSVALVKETVVISPPCTGSSSEHWKPELEBKILALEKE 1138
QY 496 KVQGSPLKPLFSSSGGLKSLGKQKRGKGDEEPEYQIHHTESPESADEQKES--- 552
Db 1139 QLOKQLQEAUTSRKAILKKAQEKERHLR-----EELQKQDDYNRLQEQDFEOKENENI 1193
QY 553 -----SASSPEEPEETTCLEKGLPEAPQDGEABEGTT----- 584
Db 1194 GDQLRQLOIQVRESIDGKLPSTDQOESCSSTPGLBEPFLPKATEQHHTQPVLESNLCPDWP 1253
QY 585 --SDGKKREGITPMAASFYKQWTPKRRVRPSBSDEKEELEKVK-SATLSSTDSVTSEMQ 641
Db 1254 SHSEDASALQGGTSVAQIKRAQL-----KEIEAEKVELKELKVSSTTSELTKKS 1300
QY 642 DEYKTVGEEQKPEPERRVDTSVSWEALICVGSKKRARKASSDDGGPRTLGGDHRA 701
Db 1301 EEFVQEQEQINKQGLBIESLKTVSHEAVHAESLQKL-----ESSQLQIAGLEHLR 1352
QY 702 EEASKKKEAGTDAVPASTODQOAGS---SSPEPAGSPSEGEVSTWESFKELVTPRKK 758
Db 1353 ELQPKLDEL---QKLISKKEEDVSYLSQSLSEKALTKIQTETIIQEEDLIKALHT---- 1405
QY 759 SKSLBEKABDSVQOLSTEI-----BPSREESVSIKKFIPGRRKRKADQKQEQATV 811
Db 1406 -QLEMQAKHEHDERIKQLOVELCEMKQKPEIGES-----RAKQOIQRKLQAAALI 1454
QY 812 EDGCPVEINEDDPNVAVPLSYNAVER-----ERKEAQCNTLPELLQAGVAVSEE 863
Db 1455 SRK---EALKENKSLQFELSAR-GTIERLTKSLADVESQVSAQNKEKDTVLGRLLQOE 1510
QY 864 LSKTLVHTVSAVIDGTAVTSVEERSPSWISASVTEPLEHTAGEAMPVVEVTEKDIIA 923
Db 1511 ERDKLITEMDRSLLENQSLSSSCC--SLKALEGLTDEKELVKEI-----ESLKSKIA 1563
QY 924 EETPVLVTQTLPEGKDAHD-----MVTSEVDTSEAVTATETSEALRTTEE----- 968
Db 1564 EST-----EWQEKHEKELQEKYEILLQSYENVSNEAERIQHVVEAVRQEKQELYGLR 1615
QY 969 VTASGAABETDMVSAVSQLTSDPTTEATPVQESGVLDTTEEROTQALQAVADK 1028
Db 1616 STEANKKETEKQLQEAQEQEEMKEMKRMKF--AKSQKQKILELEEDERLRAEVHPAGDT 1673
QY 1029 VKESQVPATQTVORTGSKALEKVEEVEDSEVLASE-----KEKDMPKGPVQAGAEH 1083
Db 1674 AKS-----CMETLLSNASMKLEELERVKMEYETLSKFFSLMSEKQSLSE-EVQD--LKH 1725
QY 1084 LAQGSBTGQATPESLEVPVTDVHDVATCQV---IKLQOLMEQAVAPESSETLTDSETN 1140
Db 1726 QIEGNVSKQANLEATEKHNDQNTVTEEGTQSIPGETEEQDLSMSRTPCTSESVPSSAKS- 1784
QY 1141 GSTPLADSDTADGTQD---ETIDSDQSKATAAVROSQVTEEEAATAQKEEPSLTPNNVP 1197
Db 1785 -ANPAVSKDFSSHDEINNYLQITDQLKRIAGLEEBEKKQNKESQTLNEK-NTLLSQIS 1842
QY 1198 AQEEHGEPEGRVLEPTQOELT-----AAAVPVLAKTEVQESGVDWLDGKVKEE 1248
Db 1843 TKD--GE-----LKMQLQEVTKQNLNNOIQIBELSRVTKLKTABEEDKDLERLMNQL 1894
```


SEQ ID NO 55868
 LENGTH: 2722
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-724-676A-55868

Query Match 5.0%; Score 400; DB 5; Length 2722;

Best Local Similarity 19.1%; Pred. No. 7.1e-15; Matches 306; Conservative 304; Mismatches 663; Indels 330; Gaps 58;

43 ADADATKLPRKNGQLSSVNGVAGDVHVOENOGGEEVVDVQGRSEEDVREKDR 102
 DB AEAEQRRLDYES-QTADNLLTEQ-IHSLISAKSKCVKI-EVLQNELDDVQ---- 867
 QY 103 VEEMANSTAVEDITKDGGEETSEII-EQIPASENNVEEMVQPAEQANDVGFKKVFK 159
 DB 868 -LQFEEQSTLLISLQSLQNKSEFEVLEGAEVRHSHSSKVEELQALSQKELEI----- 919
 QY 160 FVGPFYVKKDKNEKSDVYQLLTVKKBEGEAGVAGADHDQEPVETAVGSASKESEL 219
 DB 920 -----TKMD-----QLLEKKRIVETLQQTIEEKQOVTEISFQMTKMQVQLANEE 964
 QY 220 KOSTEKQEGTLKQEOSTEIPLQAESDQAEAEAKDEGEKQEKPTKSPESPSPVNSE 279
 DB 965 KESLQVEITLKE-----QNLISRAEAKKQVEEDNVSGLKQNYDEMSPAQGISKE 1019
 QY 280 TTSSFFKFFTHGMAWRKKTSPFKSK-----EDDLTAERKKEQEAKEYDEBEKEXT 331
 DB 1020 E-----LQHEFDLLKKEKNEQKRKLQALINRKELLQVRSLRELEILNLDESKEEIP 1072
 QY 332 EASAEQEAEDTDQARS-----ADYEKVELPLEDOVGDLA-----SSEKCAPLA 379
 DB 1073 LSETTEGEVEEDKENKESYSEKCVTSKQCEIETYLKOTISEKEVELQHIKIDLEEK--LA 1129
 QY 380 TEVPDEKMAHOEVAEVAVSTVEKTEBEOGGGGAEGGVVEGTGESLPPEKLAPOPEV 439
 DB 1130 AE-----EQFALVQNMQTLQDKTNQ-----IDLLQAEISNQMI 1165
 QY 440 POEABPAEELMKSRMCWCGGHTQLT-----LSPEEKLPRKPSGIVEVEMLSQERI 495
 DB 1166 IOK-----LITSNYDADGDSVALVETVVISPPCTSSSEHWKLELEKILAEKEKE 1218
 QY 496 KYGGSPLKLFSSSGLKLSGKKQKGRGGGDEBEKGEVQHITHESPEADBOKSES--- 552
 DB 1219 QLOKKLQELTRKALILKKAQEKERHLR-----ELKKQKQKDDNRLQEQPDEQSKENENI 1273
 QY 553 -----SASSPEEPEFTTCLKEKPLAEPQDEABEGTT----- 584
 DB 1274 GQQLQQLQIQVESIDGKLPSYDQSCSSTGPLEBRLPKATBQNHHTQVLESNLCRDPK 1333
 QY 585 --SDGEKKREGITTPWASFKRWVTPKKRVRRPESDKEELEKVK-SATLSSTDSIVSEMQ 641
 DB 1334 SHSEDSALQGGTSVAQIKQAOL-----KEIAEKVELELKVSSSTSELTYSKS 1380
 QY 642 DEVKTVGEEQKKEEPKRRVDTSVSWALLCVGSSKKRAKASSDDEGPRILGSGSHA 701
 DB 1381 BEVFOLOEOINKQGEIESLKTYSHEAEVHABSLQOKL-----ESSQLOIAGLEHLR 1432
 QY 702 BEASKDKEAGTAVPASTOEOPOAQS-----SSPEPAGSPSEGEVSTWESFRLVPRKK 758
 DB 1433 ELQPKLDEL---QKLISKEEDVSYLSQGLSEKALTLTIQIETIIEQEDLIKALHT---- 1485
 QY 759 SKSKLEKKADESSVEQLSTEI-----EPSREESVSIKIFPGRRKRADQKQEQATV 811
 DB 1486 -QLEMAKHEBRIKQLQVELCEMKQPREIGEES-----FAKQOIRKIQALALI 1534
 QY 812 EDSGPEVNEEDBNPVAVVPLSEYNAVER-----EKMEAGQNTLQQLGAVVVSSE 863
 DB 1535 SRK---EALKENKSLQEEELSLAR-GTIERLTKSLADVESQVSAONKEKDTVLGRLALQE 1590
 QY 864 LSKTLVHTVSVAVVIDTRAVTASVEERSPSMISASVTEPLEHTAGAMPVVEETEKDILA 923
 DB 1591 ERDKLITTEDRSLLENQSLSSSCG--SLKLALEGLTEDEKELVKEI-----ESIKSKSLIA 1643

QY 924 EETPVLTQTLPEKQADHD-----WTSSEVDFTSFAVTATETSEALRTEE----- 968
 DB 1644 EST-----EQEKKELOKEYEILLQSTYENVSNEHARIQHAVEAVRQKQELYGKLR 1695
 QY 969 VTEASGAETTDVMSAVSQTLDSPDTEEARPVQEVESGVLDTEEBERQIQAIIQAVADK 1028
 DB 1696 STANKKETERQLOEABEGEMEMKEMKRRKF--AKSKQKILLEENDRLRAAEVHPAGDT 1753
 QY 1029 VKESQVPATQVORTGSKALEKEYEEVEDESVLASE-----KEKDVPKQPVQDAGAEH 1083
 DB 1754 AKE-----CMETILLSNASMAKKEELERVMEYETILSKFQSLMSKSDLSSE-EVOD--LKH 1805
 QY 1084 LAQSEGTQATPESELSEPEVADVDHATCOV---IKLQMLEQAVVPESETTLDSTBN 1140
 DB 1806 QIEGVNSQKQALNLEATEKIDNOTNTERGTOSIPETEBQDSLNSSTRPTCSVSPAS- 1864
 QY 1141 GSTPLADSDTADGTQOD---ETIDSQDSKATAAARQSVTEEBATAQKEPSTLPNNVP 1197
 DB 1865 -ANPAVSKDPSSHDEINNYLQOIQLKERINGLEBEKQKNKEPQITLENK-NTLLSQIS 1922
 QY 1198 AQEHEGEEPRGDVLEPTQOELT-----AAAVPLAKTEVQGEVMDLGEKVKEE 1248
 DB 1923 TKD--GE-----LKMLOEVTKNMLNQIOEELSRVTKLKETAEERKDDLFERLMNQL 1974
 QY 1249 QEVFPHSGPNQKADVT-----YDSEYMGVAGQOEKESTEVQSLSEBEGMETDVE 1300
 DB 1975 AEL---NSIGNYQODVYDAQIKNELLESEKMLKKCVSELBEERKQOLVKEKTVSEIR 2031
 QY 1301 KEKETKPEQVSEEGEQTAPAEHGTGKPVLTLDMPSSERGALGSLGSPSLPPDQK 1360
 DB 2032 KEYLEKIOGAQKEQGNK-----SHAKLOELLKQOEYVQL---QKDCIRIQEK 2078
 QY 1361 AGCIEVOYQSLDTVTYOTAFAVEKVIETVVISFGESEPCYGAH-----LLPAKSSAT 1414
 DB 2079 ISALERVYKALEFQTESQKDLITKEN--LAQAVENHKKAQALASLAKVLLDQTQSEAA 2136
 QY 1415 GGHWTLOQAEPTVPLGPESQA--ESIPITVPAPESHTMPLQOEIAMSQREBEERKP 1472
 DB 2137 -----RVLANLKLKKELQSNKESVKSQMKQKDE---DLERRLQEAERKHLKEKNM 2185
 QY 1473 DAGPDADKESTALEKVL-----KAPEILELESKSNKIV 1507
 DB 2186 QEKDALREKVEHLEETIGEIVTLNKKQEVQQLQENLDSTV 2228

RESULT 13

US-09-724-676-55849
 Sequence 55849, Application US/09724676

GENERAL INFORMATION:

APPLICANT: Compugen LTD
 TITLE OF INVENTION: Variants of alternative splicing

FILE REFERENCE: 129181.4 Compugen
 CURRENT APPLICATION NUMBER: US/09/724,676

CURRENT FILING DATE: 2000-11-28
 NUMBER OF SEQ ID NOS: 97222

SOFTWARE: PatentIn version 3.2
 SEQ ID NO 55849

LENGTH: 2915
 TYPE: PRT

ORGANISM: Homo sapiens
 US-09-724-676-55849

Query Match

5.0%; Score 400; DB 5; Length 2915;
 Best Local Similarity 19.1%; Pred. No. 7.9e-15;
 Matches 306; Conservative 304; Mismatches 663; Indels 330; Gaps 58;

43 ADADATKLPRKNGQLSSVNGVAGDVHVOENOGGEEVVDVQGRSEEDVREKDR 102
 DB AEAEQRRLDYES-QTADNLLTEQ-IHSLISAKSKCVKI-EVLQNELDDVQ---- 787
 QY 103 VEEMANSTAVEDITKDGGEETSEII-EQIPASENNVEEMVQPAEQANDVGFKKVFK 159

Db 788 -LQFSEOSTLIRSLOLQNKSEVLEGAERVHHSKVEELSQALESKELEI----- 839
Qy 160 FVGKFTVKDKRNEKSDTVOLLTVKKDEGEGAEASVAGAGDHQEPSEVETAVGESAKSESEL 219
Db 840 -----TKMD-----QLLLEKRDVETLQOTIEEKDQVTEISFSTKMKVQNLNEE 884
Qy 220 KQSTKQEGTLKQESSTETPLQAESDQAAEEBAKDEGEKQKEPTKSPESPSPVNSE 279
Db 885 KFSGLGVEIKTLKE-----QLNLLSRABEAKEQVEEDNEVSSGLKQNYDEMSPAGQISKE 939
Qy 280 TTSSFKKFFTHGAWGRKKTSPFKSK-----EDDLETAEKKEQAEKVDDEEKEKT 331
Db 940 E-----LOHEFDLLKKEQERRKQLQALINRKELLQVRSLREBELANIKDESKKEIP 992
Qy 332 EPASEEPEAEDTDQARLS-----ADYKVELPLEDQVGDLEA-----SSEKCAPLA 379
Db 993 LSETERGEVEEDKENKEYSEKVTSKQETIYLYKQITISEKEVELQHIRKDLEBK--LA 1049
Qy 380 TEVPDEKMAHOEVAEVHVSIVTEKTEEBEOGGGGEAGGVVVEGTGSLPEKLAEPQEV 439
Db 1050 AE-----EFOALVKQMNQTLQDQTNQ-----IDLLQAEISENQAI 1085
Qy 440 POBAEPAEELMKSRMVCVSGDHTQLTD---LSPBEKTLPKHPEGIVSEVEMLSQERI 495
Db 1086 IQK-----LITSNTDASDGSVALVKEVTIVISPPCTGSEHWKPELEKILALEKEKE 1138
Qy 496 KVGSPULKLFSSSGLKJGKKQKRGGGDEPEGYQHIIHTSPESADEQKGES--- 552
Db 1139 QLQKKLQEAULTSKAILKKAQEKERHLR-----BELQKQDDYNRLQEOQDQSKENENI 1193
Qy 553 -----SASSPEEPEETTCLEKGPLEAPQDGEAEAGTT----- 584
Db 1194 GDQLRLQIQVRESIDKGLPSTDQOESCSSTPGLPEELFRKATEQHHTQPVLESNLCPDWP 1253
Qy 585 --SDGEKKREGITPWASFKKWTPKVRPSPESDKEELEKVK-SATLSTDSVTSEM 641
Db 1254 SHSDEAALOGGTSVAQIKQAOL-----KEIEAKVLELKVSTTSBELTKKS 1300
Qy 642 DEYKTVGEOKPEPKPRKRVDTSVSWEALICVSSKKKARKASSDDGGPRTLGGDSHRA 701
Db 1301 EEVFLQEQINKQGLEIETSLKTVSHEAEVHAESLQOKL-----ESSQLQIAGLEHLR 1352
Qy 702 BEASKOKEAGTDVAPASTQDQOAGS---SSPEPAGSPSEGEVSTWESFKRLVTPRKK 758
Db 1353 ELQPKLDEL---QKLISKKEEDSVLSQOLSEKAAALTKIQTIEQEDLIKALHT--- 1405
Qy 759 SKSKLEKAEDESSVEQLSTEI-----EPREESVSIKKFIPGRKKRADGQEOATV 811
Db 1406 -QLEMQAKEHDERIKQLQVLECMKKQKPEEIGES-----RAKQIQIRKLQAL 1454
Qy 812 EDGSPVEINDDPNPVPVPLSEYNAYER-----EKMEAQNTLPPOLLGAVVYSEE 863
Db 1455 SRK---EALKENKSLQELSILAR-GTIERLTKSLADVESQVSAQNKEDTVLGRLLALQE 1510
Qy 864 LSKTLVHTVSVAVIDGTRAVTSVEERSPFSWISASVTEPLEHTAGEAMPPEEVTEDKIIA 923
Db 1511 ERDKLIITEMDRSLLENQSLSSCE--SLKLALEGLTEDKEKLVKEI-----ESLKSXKIA 1563
Qy 924 EETPVLVTQTLPEGKADHD-----MVTSEVDFTSIAVATATETSEALRTEE----- 968
Db 1564 EST-----EWOEGKHEKLQKEYEILLQSYENVVSNABERIQHVVEAVRQEKQBYGLKL 1615
Qy 969 VTRASGAETTTDMVSAVSQLTDSPTDTEATPPQVESSGLVDTEEBERQTOAILQAVADK 1028
Db 1616 STEANKKETEKLOQAEQAEQMEEMKMKRP--AKSKQOKILEEEDNDRUAEVHPAGDT 1673
Qy 1029 VKESQVPATQTVORTGSKALEKVEEVEEDSEVLASE-----KEKQVMPKGPVQEGAEH 1083
Db 1674 AKE-----CMETLSSNASMKELERVKMEYETILSKFQSLMSEKDSLSE-EVQD--LKH 1725
Qy 1084 LAQGETGQATPESLEVPVETADVHDVATCOV-----IKLQQLMEQAVAPESSETITDSETN 1140
Db 1726 QIEGNVSKQANLEATEKHDNQTNTVBERTQSIIPGETEBEQSLSMSTRPTCSVPSAKS- 1784

Qy 1141 GSTPLADSDTADGTQOD---ETIDSQDSKATAAVROSQVTEEEAATAQKEPSTLPNNVP 1197
Db 1785 -ANPAVSKDPSSHDEINNYLQIDQLKERIAGLEEEKQKNKFSQTLNENEK-NTLLSQIS 1842
Qy 1198 AQEHEGEPORDVLEPTQOELT-----AAAVPVLAKTEVQEGEVDWLDGEKVEE 1248
Db 1843 TKD--GE-----LKMLOEBVTWNLNNOIQIBELGRVTKLKTAEBEKODLEERLMNQL 1894
Qy 1249 QEVFVHSGPNSSQAADVT-----YDSVMGVAGQEKESTEVQVLSLEEGEMETDVE 1300
Db 1895 AEL---NGSICNVCDVTDQAIKNELLESEMKNLKKCVSELEBKQOLVKEKTVSEIR 1951
Qy 1301 KEKRETKPEQVSEGEQETAAPBEHGYTKPVLTLDMPSSSERGKALSGSGSPSLPDQDK 1360
Db 1952 KEYLEKIQGAQKEPQGNK-----SHAKELQELLKEKQEVKQL-----QKDCIRYQEK 1998
Qy 1361 AGCIEVQVQSLDITVTCTAAEAVKVIETVVISSETGESPECVGAH-----LLPAEKSSAT 1414
Db 1999 ISALERTVKALEFVQTESQKDLITKEN--LAQAVEHRKKAQAEASFVLLDDTQSEAA 2056
Qy 1415 GGHWTLQHAEDTVPLGPESQA--ESIPITVTPAPESTLHPDLOGEISASORSESEEDKP 1472
Db 2057 -----RVLADNLKLKKELOSNEKSVSQMKQKDE-----DLERRLEQAEKHLKEKNM 2105
Qy 1473 DAGPDADGKESTAEIKVL-----KAPEPILELESKSNKIV 1507
Db 2106 QEKLDALRREKVLHEETIGETIGIQTVLNKKDKKEVQOQENLDSTV 2148
RESULT 14
US-09-724-676-55850
; Sequence 55850, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 55850
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-55850

Query Match 5.0%; Score 400; DB 5; Length 2915;
Best Local Similarity 19.1%; Pred. No. 7.9e-15;
Matches 306; Conservative 304; Mismatches 663; Indels 330; Gaps 58;

Qy 43 ADADPATKLPQKNGQLSSVNGVAEGDQVHVQENQEGQEEVVDVEDVQGESEDVREKDR 102
Db 738 AEAEORRLDYES-QTAHDNLLTEQ--IHSLSIEAKSKDVKI---EVLQNELDDVQ--- 787
Qy 103 VEEMAANSTAVEDITKDGQOETSEII---EQIPASENNVEEMVQPAESQANDVGFKKVFK 159
Db 788 -LQFSEOSTLIRSLOLQNKSEVLEGAERVHHSKVEELSQALESKELEI----- 839
Qy 160 FVGKFTVKDKRNEKSDTVOLLTVKKDEGEGAEASVAGAGDHQEPSEVETAVGESAKSESEL 219
Db 840 -----TKMD-----QLLLEKRDVETLQOTIEEKDQVTEISFSTKMKVQNLNEE 884
Qy 220 KQSTKQEGTLKQESSTETPLQAESDQAAEEBAKDEGEKQKEPTKSPESPSPVNSE 279
Db 885 KFSGLGVEIKTLKE-----QLNLLSRABEAKEQVEEDNEVSSGLKQNYDEMSPAGQISKE 939
Qy 280 TTSSFKKFFTHGAWGRKKTSPFKSK-----EDDLETAEKKEQAEKVDDEEKEKT 331
Db 940 E-----LOHEFDLLKKEQERRKQLQALINRKELLQVRSLREBELANIKDESKKEIP 992
Qy 332 EPASEEPEAEDTDQARLS-----ADYKVELPLEDQVGDLEA-----SSEKCAPLA 379

```

Db 993 LSETERGEVEDEKENKEYSEKCVTSKCOEIELYKOTISEKEVELOHTRKDLK--LA 1049
Qy 380 TEVFERKEAEHQEVVAEYHVSTVEKTEBEOGGGEAEGVVETGESLPREKLAPEPEV 439
Db 1050 AE-----EOPQALVKQNMOTLODKTNO-----IDLQAEISENOAI 1085
Qy 440 POEAPAEELKMSRMCVSGGDHQTLD-----LSPEKTLPKHPEGIVSEVEMLSOBRI 495
Db 1086 IQK-----LITSNTPDASDGSVALVKEVTIVISPPCTGSSEHMKPELEKILALEKEKE 1138
Qy 496 KVQSGPLKLPLSSSGIKLISGKKQKRGKGDEEPGEYOHITHESPSADBOKSES--- 552
Db 1139 QLOKKLOALTSRKAILKKAQEKERHLR---EELKQKODYVRLQEPDQOSKENENI 1193
Qy 553 -----SASSPEPEBETTCLKGPLEAPDQGAEBGTT----- 584
Db 1194 GDQLRQLOIQVRESIDGKLPTDQOESCSPTGLEPPLFKATEOHHTOPVLESNLCPPMP 1253
Qy 585 --SDGEKREGETTPWASPRKMWTPPKKVRPRSESDKEELEKVK--SATLSTDSPTVSEMQ 641
Db 1254 SHSEBASALQGGTSVAQIKAOI-----KEIEAEVLELELKVSSTTSSELTKKS 1300
Qy 642 DEVTKVGEOQPEEBPKRVDTSVSMELICVSGSKRKARAKASSDDEGGPTLGGDSHRA 701
Db 1301 EBFQLOBOINKQCLEISLKTIVSHEAEVHAESLQOKL-----ESSOLOIAGLEHLR 1352
Qy 702 EBAASKDEAGTDVAVASTQEOBQOAGS---SSPEPAGSPEBEGVSTWESFRVLTPRKK 758
Db 1353 ELQPLRDEL---QKLISSKEBVSYLSGOLSEKKAALTKIQEIIIEQEDLIKALHT--- 1405
Qy 759 SKSKLEBAEDSSVQOLSTEL-----EPSREBSWISIKRTPGRKKRAGKQBOAT 811
Db 1406 -QLEQOAKHEHERIQLOVELCEMKQKPELIGES-----RAKQOIORKQOALI 1454
Qy 812 EDSGVEINEDDPNPAVPAVPLSEYNAVER-----EKMEAQGTTELPLQLAGAVVSE 863
Db 1455 SRK---EALKKESKIQEBELSLAR-GTIERLTSLADVEQVQAOKKQDVTGRLALIOE 1510
Qy 864 LSKTLVHTVSAVAVIDGTRAVTSVEERSPWISASVTEPLEHTAGAMPVEEVEKDIITA 923
Db 1511 ERDKLITEMDSILLENQSLSSCE--SLKLALEGLTEDKELVKEL-----ESLSSKIA 1563
Qy 924 EETPVLTTLEPGKAND-----MTSEVDFTSEAVTARETSBALRTEE----- 968
Db 1564 EST-----EMOKKHELOKEYEYILLQSYENVSNEARIGHVAVAEQOEBLYGKL 1615
Qy 969 VTEASGAETTDMSAVASOLTPSTTEATPVQVESGVLDTEEBERTOAILQAVADK 1028
Db 1616 STEANKKTEKOLQAEQOEMEMKEMKRF--AKSKQOKILELEENDRILRAEVRPADDT 1673
Qy 1029 VKESQVPAOTVORTGSKALEKEVEBEDSEVLASE-----KEKDVMKGPQVQAGAEH 1083
Db 1674 AKE-----CMETLLSSNMSKMEKELERVKMEYETLSKKFQSLMSEKOSLSE--EVQD--LKH 1725
Qy 1084 LAQSGSETQATPESLEVEPVTAADVHVATQV---IKLOQLEQAVAPSSSTLLDSEIN 1140
Db 1726 QIEGVNSQOAMLEATEKHDNQTNTVEEGTQSIPEGTEEBODSLMSSTRPTCSSEVSAXS- 1784
Qy 1141 GSTPLADBDTADGTQOD---ETIDSQDSKATAAVASQVTEBEAATAQKEEBSSTLPNNVP 1197
Db 1785 -ANPVSXKDFSSHDEINNYLOQIDQKERIAGLEBEKQKNKEFSOTLENEK--NTLLSQIS 1842
Qy 1198 AOEBHGEPEPRDVLEPTQOELT-----AAAVPVLAKTEVGOEGEDVMDJGKVKAE 1248
Db 1843 TKD--GE-----LKMLOGEVTKMMLNLOQIOBELSRVTKLKTAEHEKDDLEBKLMNQ 1894
Qy 1249 QEVFVHSGPNSQKADVT-----YDSEVWGAVAGQEKESTEVQSLSEBGEWETDVE 1300
Db 1895 AEL---NGSIGNYCODVTDQAIKNELISEMKNLKKCVSELEEEKQOOLVKEKTVKESIR 1951
Qy 1301 KEKRETKBEQVSEBGOETAAPEHNGTYGKPVLLTDMPSSENGKXLAGSLGSPSLPDDOX 1360
Db 1952 KEYLEKIQAOKEPGRK-----SHAKELQELLKEKQOEVKOL---QKDCIRYQEK 1998

```

```

Qy 1361 AGCIEVOYQSLDTTVOYTAENAEVKEVIEVVISETGESPECYGAH-----LLPAKKSAT 1414
Db 1999 ISALERTVKALEFVQTESQKLETTKEN--LAQAEHKKQAQELASPRVLLDITOSERA 2056
Qy 1415 GGHMTLOGAEDVPLPGPSQA--ESIPITVPAPESTLHPDLOEISASQRESRESEEDKP 1472
Db 2057 -----RVLADNLTKLELOJNKESVKSQMKQDE-----DLERLEQAEKHLKEKQNM 2105
Qy 1473 DAGPDADGKESTALEKVL-----KAPELILESKSNKIV 1507
Db 2106 QEKDALRREKHYHEETIGEIOVTLNKKKREVOQLQEWLSDTV 2148

RESULT 15
US-09-724-676A-55849
; Sequence 55849, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Comugen LTD
; TITLE OR INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Comugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 55849
; LENGTH: 2915
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-724-676A-55849

Query Match      5.0%, Score 400, DB 5, Length 2915;
Best Local Similarity 19.1%, Pred. No. 7.9e-15;
Matches 306; Conservative 304; Mismatches 663; Indels 330; Gaps 58;

Qy 43 ADADATVTLPGKNGQLSGVNGVAGQDVHVOEENQOGEVEEVDEVDQORESEPDREKDR 102
Db 738 AEAERORRLDYES--OTARDNLITLDO--IHSLSIEAKSKADVXI---EVLQNLBDVQ--- 787
Qy 103 VEEMANSTAVADITTKQOEBTSEIIT---EQIPASENNVEEMVOPAESQANDVGFKKVF 159
Db 788 -LQPEBQSTLIRLSQOLQNKSESEVLBGAERVHRHSKVEELSQLSQKLEI----- 839
Qy 160 FVGKFTYKKDKNEKSDTVQLLTVYKQGBGAENASVGAQHOEBSVETAVGASAKSESL 219
Db 840 -----TKMD-----QLLLEKKRDVETLQOTIEEKDQOVTEISFSMTEKMWQLNBE 884
Qy 220 KQSTFKQSGTLKQOBSSSTEIPLQAESDQAAEEAKDEBEEKQEKPTSPSPSPVNSE 279
Db 885 KPSLQVEIKTLKE---QNLNLSRAEBAKKEQVBEDEVSSGLKQNYDEMSPAQISKE 939
Qy 280 TTSSFKEPFTGAWGRKKTSPFKSK-----EDLETAEKREKQEAKEVDEBEKEKT 331
Db 940 E-----LQHEFDLKKENQQRKXLOALINKKELLQVRSLEELANLKDESKKILP 992
Qy 332 EPASEQEPADTDQARIS---ADVEKVELPLEDYQGDLEA-----SSEKCAPIA 379
Db 993 LSETERGEVEDEKENKEYSEKCVTSKCOEIRIYLKOTISEKEVELOHTRKDLK--LA 1049
Qy 380 TEVFERKEAEHQEVVAEYHVSTVEKTEBEOGGGEAEGVVETGESLPREKLAPEPEV 439
Db 1050 AE-----EOPQALVKQNMOTLODKTNO-----IDLQAEISENOAI 1085
Qy 440 POEAPAEELKMSRMCVSGGDHQTLD-----LSPEKTLPKHPEGIVSEVEMLSOBRI 495
Db 1086 IQK-----LITSNTPDASDGSVALVKEVTIVISPPCTGSSEHMKPELEKILALEKEKE 1138
Qy 496 KVQSGPLKLPLSSSGIKLISGKKQKRGKGDEEPGEYOHITHESPSADBOKSES--- 552
Db 1139 QLOKKLOALTSRKAILKKAQEKERHLR---EELKQKODYVRLQEPDQOSKENENI 1193
Qy 553 -----SASSPEPEBETTCLKGPLEAPDQGAEBGTT----- 584

```


Db 1194 GDQLRLOIQIVRESIDGKLPSTDOOESCSSTPGLLEPLFKATEQHHTQPVLESNLCPDWP 1253
QY 585 --SDGEKKRGITPWSFKKXKVTWKVRPSSESDKEELEKVK-SATLSSTSTVSEM 641
Db 1254 SHSEDASALQGGTSVAQIKQAQL-----KEIEAKVELELKVSTTSELTKS 1300
QY 642 DEVKTVGEEQKPEEPKRRVDTSVSWEALICVGSKKKARKASSDDDEGGPTLGGDSHRA 701
Db 1301 EEVFOQEQINKOGLETESLKTVSHEAVHAESLQOKL-----ESSOLQIAGLEHLR 1352
QY 702 EEASKDKEAGTDAVPASTQDQAQGS---SSPEPAGSPSEGEVSTWESFKRLVTPRKK 758
Db 1353 ELQPKLDEL---QKLSKKEEDVYLSGQLSEKAALTKIQTIIIEQEDLIKALHT---- 1405
QY 759 SKSKLEKAEDSSVQELSTEI-----EPSREESVWSIKKPIPRKKRADGKQEOATV 811
Db 1406 -QLEMQAKEHDERIKOLQVELCEMKQKPEIGBES-----RAKQOIQRKLOAALI 1454
QY 812 EDSPVNEINDDPNVPAVPLSEYNAVER-----EKMEAOGNTELPOLLGNAVYSEE 863
Db 1455 SRK---EALKENKSLQBELSLAR-GTIERLTKSLADVESQVSAQNKEKOTVIGRLALLQE 1510
QY 864 LSKTLVHTSVAVIDGTRAVTSVEERSPSWISASVTEPLEHTAGAEAMPPVEEVTEKDIIA 923
Db 1511 ERDKLITEMDRSLLNQSLSSCE--SLKALGLTEDEKELVKEI-----ESLKSKKIA 1563
QY 924 EETPVLTTQTLPEGKDAHD-----MVTSEVDPTSAAVTATSETSEALRTEE----- 968
Db 1564 EST-----EWOEKHELQKEVEILLQSVNVNSNEARIQHVVAVRQEKQELYGKLR 1615
QY 969 VTASGAEETDDMVSAVSQLTSDPTTEATPVQEVESGLDTEEBERQTAILOAVADK 1028
Db 1616 STEANKKETEKQAEAEQEMEEMKEKMKRP--AKSKOOKILELEEEENDRLRAEVHPAGDT 1673
QY 1029 VKESQVPAQTQVORTGSKALEKVEEVEDSEVLASE-----KEQVMPKGPVQEAAGAH 1083
Db 1674 AKE-----CMETLLSSNASMKEELERVMEYETLSKKFQSLMSEKDSLSE-EVQD--LKH 1725
QY 1084 LAQGETGQATPESLEVEPVADVDHVATQCV---IKLQQLMEQAVAPESSETLTDSETN 1140
Db 1726 QIEGNVSKQANLEATEKHNDQNTVTEBGTQSI PGTEEBEQDQSLSMSTRPTCSSESVP SAKS- 1784
QY 1141 GSTPLADSDTADGTQDP---ETDSDQSKATAAVRQSVQVTEEBEATAQKEEPSTLPNNVP 1197
Db 1785 -ANPAVSKDFSSHDEINNYLQOIDQLKRIAGLEEERQKNEKFSQTLNEK-NTLLSQIS 1842
QY 1198 AQEBHGEPGRDVLPTQOELT-----AAAVPVLAKTEVQGEQGEVDWLDGEKVKEE 1248
Db 1843 TKD--GE-----LKMLOEVTWNLNQIOBELSRVTKLKTABEEKODLEERLMNQL 1894
QY 1249 QEVFVHSGPNSOKAADVT-----YDSEVMGVAGQOEKESTEVSQSLSEEGEMETDVE 1300
Db 1895 AEL---NGSIGNYCQDVTDAIKNELLESEMKNLKKVCVSELEBEKQOLVKEKTKVESEIR 1951
QY 1301 KEKRETPQEVQSEGEQETAPEHEGTYGPVLTLDMPSSBERGKALGSLGSPSLPDQDK 1360
Db 1952 KEYLEKTOGAQKEPGNK-----SHAKELQELLKQOEKVQKQ-----QKDCIRYQEK 1998
QY 1361 AGCIEVOVQSLDTTQVTAABAVKVIETVWISETGESPECVGAH-----LLPAEKSSAT 1414
Db 1999 ISALERTVKALEFVQTESQKDLITKEN--LAQAVEHRKKAQELASFVLLDDTQSEAA 2056
QY 1415 GGHWTLQHAEDTVPLGPESAQ--ESIPITVPAPESTLHPDLOGEISASQERSEEDKP 1472
Db 2057 -----RVLADNLKLKKELOSNKESVSKQMKDE-----DLERRLEQABEKLKEKKNM 2105
QY 1473 DAGPDADGKESTAIEKVL-----KABPEILELESKNKIV 1507
Db 2106 QEKLDALRRREKRVHLEETIGETIQVTLNKKQKEVQOQLENLDSTV 2148

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 13, 2002, 00:28:41 ; Search time 6461 Seconds
(without alignments)
6210.712 Million cell updates/sec

Title: US-09-902-432-4
Perfect score: 8073
Sequence: 1 MGNSTGSRSEQPGASDT.....AWAQRKCLPRLQKAPVSK 1596

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 24791104 seqs, 12571243825 residues

Total number of hits satisfying chosen parameters: 49582208

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-O=/cgn2_1/USPTO_spool/US0902432/runat_12122002_141453_25861/app_query.fasta_1.1735
-DB=Pending Patents_NA_Main -OPMT=fastap -SUFFIX=nrpm -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0902432 @CGN 1 1 5452 @runat_12122002_141453_25861 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Pending Patents_NA_Main:
1: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq.*
2: /cgn2_6/ptodata/2/pna/US06_COMB.seq.*
3: /cgn2_6/ptodata/2/pna/US07_COMB.seq.*
4: /cgn2_6/ptodata/2/pna/US08_COMB.seq.*
5: /cgn2_6/ptodata/2/pna/US081_COMB.seq.*
6: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
7: /cgn2_6/ptodata/2/pna/US083_COMB.seq.*
8: /cgn2_6/ptodata/2/pna/US084_COMB.seq.*
9: /cgn2_6/ptodata/2/pna/US085_COMB.seq.*
10: /cgn2_6/ptodata/2/pna/US086_COMB.seq.*
11: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
12: /cgn2_6/ptodata/2/pna/US088_COMB.seq.*
13: /cgn2_6/ptodata/2/pna/US089_COMB.seq.*
14: /cgn2_6/ptodata/2/pna/US090_COMB.seq.*
15: /cgn2_6/ptodata/2/pna/US091_COMB.seq.*
16: /cgn2_6/ptodata/2/pna/US092_COMB.seq.*
17: /cgn2_6/ptodata/2/pna/US093_COMB.seq.*
18: /cgn2_6/ptodata/2/pna/US094_COMB.seq.*
19: /cgn2_6/ptodata/2/pna/US095A_COMB.seq.*
20: /cgn2_6/ptodata/2/pna/US095B_COMB.seq.*
21: /cgn2_6/ptodata/2/pna/US095C_COMB.seq.*
22: /cgn2_6/ptodata/2/pna/US095D_COMB.seq.*
23: /cgn2_6/ptodata/2/pna/US096A_COMB.seq.*
24: /cgn2_6/ptodata/2/pna/US096B_COMB.seq.*
25: /cgn2_6/ptodata/2/pna/US096C_COMB.seq.*
26: /cgn2_6/ptodata/2/pna/US096D_COMB.seq.*
27: /cgn2_6/ptodata/2/pna/US096E_COMB.seq.*
28: /cgn2_6/ptodata/2/pna/US097A_COMB.seq.*

29: /cgn2_6/ptodata/2/pna/US097B_COMB.seq.*
30: /cgn2_6/ptodata/2/pna/US097C_COMB.seq.*
31: /cgn2_6/ptodata/2/pna/US098A_COMB.seq.*
32: /cgn2_6/ptodata/2/pna/US098B_COMB.seq.*
33: /cgn2_6/ptodata/2/pna/US098C_COMB.seq.*
34: /cgn2_6/ptodata/2/pna/US099A_COMB.seq.*
35: /cgn2_6/ptodata/2/pna/US099B_COMB.seq.*
36: /cgn2_6/ptodata/2/pna/US099C_COMB.seq.*
37: /cgn2_6/ptodata/2/pna/US099D_COMB.seq.*
38: /cgn2_6/ptodata/2/pna/US100A_COMB.seq.*
39: /cgn2_6/ptodata/2/pna/US100B_COMB.seq.*
40: /cgn2_6/ptodata/2/pna/US101A_COMB.seq.*
41: /cgn2_6/ptodata/2/pna/US101B_COMB.seq.*
42: /cgn2_6/ptodata/2/pna/US102A_COMB.seq.*
43: /cgn2_6/ptodata/2/pna/US102B_COMB.seq.*
44: /cgn2_6/ptodata/2/pna/US6000_COMB.seq.*
45: /cgn2_6/ptodata/2/pna/US6001_COMB.seq.*
46: /cgn2_6/ptodata/2/pna/US6002_COMB.seq.*
47: /cgn2_6/ptodata/2/pna/US6003_COMB.seq.*
48: /cgn2_6/ptodata/2/pna/US6004_COMB.seq.*
49: /cgn2_6/ptodata/2/pna/US6005_COMB.seq.*
50: /cgn2_6/ptodata/2/pna/US6006_COMB.seq.*
51: /cgn2_6/ptodata/2/pna/US6007_COMB.seq.*
52: /cgn2_6/ptodata/2/pna/US6008_COMB.seq.*
53: /cgn2_6/ptodata/2/pna/US6009_COMB.seq.*
54: /cgn2_6/ptodata/2/pna/US6010_COMB.seq.*
55: /cgn2_6/ptodata/2/pna/US6011_COMB.seq.*
56: /cgn2_6/ptodata/2/pna/US6012_COMB.seq.*
57: /cgn2_6/ptodata/2/pna/US6013_COMB.seq.*
58: /cgn2_6/ptodata/2/pna/US6014_COMB.seq.*
59: /cgn2_6/ptodata/2/pna/US6015_COMB.seq.*
60: /cgn2_6/ptodata/2/pna/US6016_COMB.seq.*
61: /cgn2_6/ptodata/2/pna/US6017_COMB.seq.*
62: /cgn2_6/ptodata/2/pna/US6018_COMB.seq.*
63: /cgn2_6/ptodata/2/pna/US6019_COMB.seq.*
64: /cgn2_6/ptodata/2/pna/US6020_COMB.seq.*
65: /cgn2_6/ptodata/2/pna/US6021_COMB.seq.*
66: /cgn2_6/ptodata/2/pna/US6022_COMB.seq.*
67: /cgn2_6/ptodata/2/pna/US6023_COMB.seq.*
68: /cgn2_6/ptodata/2/pna/US6024_COMB.seq.*
69: /cgn2_6/ptodata/2/pna/US6025_COMB.seq.*
70: /cgn2_6/ptodata/2/pna/US6026_COMB.seq.*
71: /cgn2_6/ptodata/2/pna/US6027_COMB.seq.*
72: /cgn2_6/ptodata/2/pna/US6028_COMB.seq.*
73: /cgn2_6/ptodata/2/pna/US6029_COMB.seq.*
74: /cgn2_6/ptodata/2/pna/US6030_COMB.seq.*
75: /cgn2_6/ptodata/2/pna/US6031_COMB.seq.*
76: /cgn2_6/ptodata/2/pna/US6032_COMB.seq.*
77: /cgn2_6/ptodata/2/pna/US6033_COMB.seq.*
78: /cgn2_6/ptodata/2/pna/US6034_COMB.seq.*
79: /cgn2_6/ptodata/2/pna/US6035_COMB.seq.*
80: /cgn2_6/ptodata/2/pna/US6036_COMB.seq.*
81: /cgn2_6/ptodata/2/pna/US6037_COMB.seq.*
82: /cgn2_6/ptodata/2/pna/US6038_COMB.seq.*
83: /cgn2_6/ptodata/2/pna/US6039_COMB.seq.*
84: /cgn2_6/ptodata/2/pna/US6040_COMB.seq.*
85: /cgn2_6/ptodata/2/pna/US6041_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8073	100.0	5200	13	US-08-978-277A-3
2	8073	100.0	6160	1	PCT-US97-06830-3
3	8073	100.0	6160	10	US-08-665-401-3
4	8073	100.0	6160	34	US-03-902-432-3
5	8030	99.5	5236	1	PCT-US02-16173A-2063
6	8030	99.5	5236	41	US-10-191-803-279

Sequence 3, Appli
Sequence 3, Appli
Sequence 3, Appli
Sequence 2063, Ap
Sequence 279, App

Db 661 CAATCCACAGAGCAAGAGCCCTGAAGCAAGAACAGAGCAGCACAGAAATCCCC 720
 Qy 241 LeuGlnAlaGluSerAspGlnAlaAlaGluGluAlaLysAspGluGlyGluGluLys 260
 Db 721 CTTCAAGCCGAATCTGATCAAGCGGCTGAGGAAGAACCCAAAGATGAAGAGAGAAAA 780
 Qy 261 GluGluLysGluProThrLysSerProGluSerProSerSerProValAenSerGluThr 280
 Db 781 CAAGAGAAAGAGGCCCAAGTCCCCAGAAATCCCGAGCAGCCAGTCAACAGTGAACA 840
 Qy 281 ThrSerSerPheLysLysPheThrHisGlyTyrAlaGlyTyrArgLysLysThrSer 300
 Db 841 ACATCTCTTCAAGAGTTCTTCACTACGGTTGGCCCGCTGGCCCAAGAGACAGC 900
 Qy 301 PheLysLysSerLysGluAspAspLeuGluThrAlaGluLysArgLysGluGlnGluAla 320
 Db 901 TTCAAGAAATCAAAAGAGGATGATCTGAAACTGCCGAGAAGAGAAAGAGCAAGGCA 960
 Qy 321 GluLysValAspGluGluGluLysGluLysThrGluProAlaSerGluGluGlnGluPro 340
 Db 961 GAAAAAGTAGACGAGGAAGAAAGGAAAGACAGAGCCCTCGGAGGAGCAGGAGCG 1020
 Qy 341 AlaGluAspThrAspGlnAlaArgLeuSerAlaAspTyrGluLysValGluLeuProLeu 360
 Db 1021 GCAGAGACACACAGCCAGGCCAGGTGTGACGACACTACAGAGAGGTGGAGCTGCCTTTG 1080
 Qy 361 GluAspGlnValGlyAspLeuGluAlaSerSerGluGluLysCysAlaProLeuAlaThr 380
 Db 1081 GAAGACCAGGTGTGTGACCTGGAGGCATCGTCAGAGAGAGAGTGTGCTCTTTGGCAAG 1140
 Qy 381 GluValPheAspGluLysMetGluAlaHisGlnGluValValAlaGluValHisValSer 400
 Db 1141 GAAGTGTGTTGATGAGAAGATGGAAGCCACCAAGAAATGTTGTGCAGAGGTCCACGTGAGC 1200
 Qy 401 ThrValGluLysThrGluGluGlnGlyGlyGlyGlyGluAlaGluGlyGlyValVal 420
 Db 1201 ACCGTGGAAGACAGAGGAGGAGCAGGAGAGGAGGAGGAGGCTGAAGGGGGCGTGGTG 1260
 Qy 421 ValGluGlyThrGlyGluSerLeuProProGluLysLeuAlaGluProGlnGluValPro 440
 Db 1261 GTAGAGGAACAGAGAAATCCTTTGCCCTGAGAAACTGCTGAGCCCAAGAGGTCCCC 1320
 Qy 441 GlnGluAlaGluProAlaGluGluLeuMetLysSerArgGluMetCysValSerGlyGly 460
 Db 1321 CAGGAAGCTGAGCGTGTGAGGAGCTGATGAAGAGCAGAGAGATGTGTCTCTGGAGGA 1380
 Qy 461 AspHisThrGlnLeuThrAspLeuSerProGluGluLysThrLeuProLysHisProGlu 480
 Db 1381 GACCACACTCACTGACAGACCTAAGTCTTGAAGAGAGAGCGTGGCCCAACACCCAGAA 1440
 Qy 481 GlyIleValSerGluValGluMetLeuSerSerGlnGluArgIleLysValGlnGlySer 500
 Db 1441 GGCATTGTAGTGGGTGGAGATGCTGTCTCTCAGAAAGAAATCAAGGTACAGGAGAGT 1500
 Qy 501 ProLeuLysLysLeuPheSerSerSerGlyLeuLysLysLeuSerGlyLysLysGlnLys 520
 Db 1501 CCCTTGAAGAAACTCTTTCAGTAGCTCAGGCTTAAAGAAGCTGTCTGGGAAGAGCAGAA 1560
 Qy 521 GlyLysArgGlyGlyGlyAspGluGluProGlyGluThrGlnHisIleHisThrGlu 540
 Db 1561 GGGAAACGAGAGGTGGGGAGAGCAAGAGCCTGGAGAAATACCAACATTCACCCGAA 1620
 Qy 541 SerProGluSerAlaAspGluGlnLysGlyLysSerAlaSerSerProGluGluPro 560
 Db 1621 TCCCCAGAGAGTGTGATGACAGAGGAGAGAGCTTCTGCTGCTCCCCCGAGAGCCT 1680
 Qy 561 GluGluThrThrCysLeuGluLysGlyProLeuGluAlaProGlnAspGlyGluAlaGlu 580
 Db 1681 GAGGAGACCACGTGTCTGGAGAAAGGCGCGCTGGAACACCCCGAGGATGGGAAGCTGAG 1740
 Qy 581 GluGlyThrThrSerAspGlyGluLysLysArgGluGlyIleThrProThrAlaSerPhe 600
 Db 1741 GAAGGAACTACTTCCGATGGAGAGAGAGAGAGGAGATCACTCCCTGGGCATCCTTC 1800

Qy 601 LysLysMetValThrProLysLysArgValArgArgProSerGluSerAspLysGluGlu 620
 Db 1801 AAAAAGATGGTGACACCCCAAGAAACGGGTCCGAAGACCTTCTGAGAGTGACAAGAGGAA 1860
 Qy 621 GluLeuGluLysValLysSerAlaThrLeuSerSerThrAspSerThrValSerGluMet 640
 Db 1861 GAGCTGGAGAAAGGTCAAGAGCGCCACTTGTCTCCACTGATAGCACAGTGTGAGAAATG 1920
 Qy 641 GlnAspGluValLysThrValGlyGluGlnLysProGluGluProLysArgVal 660
 Db 1921 CAAGATGAAGTCAAAACTGTGTGTGAGGAACAAAAGCCAGAGAAACCAAGCGTAGGGTG 1980
 Qy 661 AspThrSerValSerTrpGluAlaLeuIleCysValGlySerSerLysLysArgAlaArg 680
 Db 1981 GATACCTTCAGTGTCTTGGGAAGCACCTGATTTGTGTCCGATCATCCCAAGAGAGCAAGG 2040
 Qy 681 LysAlaSerSerSerAspAspGluGlyGlyProArgThrLeuGlyGlyAspSerHisArg 700
 Db 2041 AAGGCATCTCTTCAGATGATGAAGAGGGGCCAAGGACACTGGGAGGGGACAGTCAAGA 2100
 Qy 701 AlaGluGluAlaSerLysAspLysGluAlaGlyThrAspAlaValProAlaSerThrGln 720
 Db 2101 GCAGAGAGGCCAGCAAGCAAAAGAACCGCGNACAGACGCTGTTCTGCCAGACCCAG 2160
 Qy 721 GluGlnAspGlnAlaGlnGlySerSerSerProGluProAlaGlySerProSerGluGly 740
 Db 2161 GAGCAGGACCAAGCGCAAGAAAGTTCCTCACCGAGCCAGCGGAAGCCCTTCCGAAGGG 2220
 Qy 741 GluGlyValSerThrTrpGluSerPheLysArgLeuValThrProArgLysLysSerLys 760
 Db 2221 GAAGGTGTCTCCACTTGGGAGTCAITTTAAAGATTAGTCACTCCAAGAAAAAATCCAAG 2280
 Qy 761 SerLysLeuGluGluLysAlaGluAspSerSerValGluGlnLeuSerThrGluIleGlu 780
 Db 2281 TCAAAACCTGGAAGAAAGACCGGAAGACTCTAGTGTAGACAGATTGCTCACTGAGATCGAA 2340
 Qy 781 ProSerArgGluGluSerTrpValSerIleLysLysPheIleProGlyArgArgLysLys 800
 Db 2341 CCGAGTAGAGAAAGAACTTTGGGTTCATTAAGAAATTCATCCCGGCGCGNAGAAA 2400
 Qy 801 ArgAlaAspGlyLysGlnGluGlnAlaThrValGluAspSerGlyProValGluIleAen 820
 Db 2401 AGCGAGAGCGGGAAGCAAGAACCAAGCACTGTGGAAGACTCAGGGCCAGTGGAGATAAT 2460
 Qy 821 GluAspAspProAsnValProAlaValValProLeuSerGluTyrAsnAlaValGluArg 840
 Db 2461 GAGGACGACCTTAATGTCCAGCCGTCGTCTCTGAGTATATATGACGTGGAGAGG 2520
 Qy 841 GluLysMetGluAlaGlnGlyAsnThrGluLeuProGlnLeuLeuGlyAlaValTyrVal 860
 Db 2521 GAGAAGATGGAAGCCAGGGGAATACGGAGCTGCCAGCTGCTGGGGCTGTGTACGTG 2580
 Qy 861 SerGluGluLeuSerLysThrLeuValHisThrValSerValAlaValIleAspGlyThr 880
 Db 2581 TCCGAGGAGCTCAGTAAAGACTCTGTGTCCCACTGTGAGTGTGTCAGTCAATTGATGGACC 2640
 Qy 881 ArgAlaValThrSerValGluArgSerProSerTrpIleSerAlaSerValThrGlu 900
 Db 2641 AGGGCAGTCACCAAGTTCGAAGAGCGGTCTCTCTCTGTGATATATCCCGTCCCGTAACAGAA 2700
 Qy 901 ProLeuGluHisThrAlaGlyGluAlaMetProProValGluGluValThrGluLysAsp 920
 Db 2701 CCTCTTGAACACACAGCGGGAGAGCCATGCCACCTGTTGAAGAGGTCACTGAAAAAGAC 2760
 Qy 921 IleIleAlaGluGluThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHis 940
 Db 2761 ATCATTGCAGAGAAACTCTCTGTGTCTCACCAGACGTTACCAGAGGTTAAAGATGCCAT 2820
 Qy 941 AspAspMetValThrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThr 960
 Db 2821 GACGACATGGTCAACAGTGAAGTGGATTTCACTCAAGAGCTGTGACAGCACAGAGACC 2880

SOFTWARE: FastSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/06830
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/665,401
FILING DATE: 18-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Richard S.
REGISTRATION NUMBER: 26,154
REFERENCE/DOCKET NUMBER: A30558 - 165/34008
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-408-2558
TELEFAX: 212-765-2519
TELEX:

Alignment Scores:		
Pred. No.:	0	Length: 6160
Score:	8073.00	Matches: 1596
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	100.00%	Indels: 0
DB:	1	Gaps: 0

US-09-902-432-4 (1-1596) x PCT-US97-06830-3 (1-6160)

QY	1	MetGlyAlaGlySerSerThrGluGlnArgSerProGluGlnProAlaGlySerAspThr	20
DB	33	ATGGCGCAGGCGAGTTCACCCAGCAGCGAGCCCGAGCAGCGCGGAGCGACG	92
QY	21	ProSerGluLeuValLeuSerGlyHisGlyProAlaGluAlaSerGlyAlaAlaGly	40
DB	93	CCGAGCGAGCTGGTGCTCAGTGGCCATGGGCCCGCAGCTGAAGCTCGGAGCAGCTGGA	152
QY	41	AspProAlaAspAlaAspProAlaThrLysLeuProGlnLysAsnGlyGlnLeuSerSer	60
DB	153	GACCCCGCGAGCGGACCCGCCACCACCAAGTCCACAGAAAGATGGCCAGCTGCTCTT	212
QY	61	ValAsnGlyValAlaGluGlnGlyAspValHisValGlnGluGluAsnGlnGluGlyGln	80
DB	213	GTCAACGGCGGTAGCTGAACAAGGAGATGTCATGTCCAAGAGGAAAAACAGGAGGGCGAG	272
QY	81	GluGluGluValValAspGluAspValGlyGlnArgGluSerGluAspValArgGluLys	100
DB	273	GAGGAAGAAGTGGTTGATGAGGATGTTGGACGCGAGAGTCAGAAGATGTGAGAGAAAA	332
QY	101	AspArgValGluGluMetAlaAlaAsnSerThrAlaValGluAspIleThrLysAspGly	120
DB	333	GACCGAGTTGAAGAAATGGCGCCAACTCCACAGCTGTTGAAGATATCACAAAGGATGG	392
QY	121	GlnGluGluThrSerGluIleIleGluGlnIleProAlaSerGluAsnAsnValGluGlu	140
DB	393	CAGGAGGAGACATCAGAATAATTAACACAGATCCCTGCTTCAGAAAAACAATGTGGAAGAA	452
QY	141	MetValGlnProAlaGluSerGlnAlaAsnAspValGlyPheLysLysValPheLysPhe	160
DB	453	ATGGTACAGCCTGCTGAGTCCAGGCTTAATGATGTTGGCTTCAAGAAAGTATTTAAATTT	512
QY	161	ValGlyPheLysPheThrValLysLysAspLysAsnGluLysSerAspThrValGlnLeu	180

D	b	513	GTTGGTTTTAAATTACCGGTGAAGAGGATAAAAAATGAAAAGTCAGATACTGTCTCCAACTA	572
Q	y	181	LeuThrValLysLysAspGluGlyGluGlyAlaGluAlaSerValGlyAlaGlyAspHis	200
D	b	573	CTCAGTGTCAAGAAGGATGAAGGCGAAGGGCAGAAAGCCTCTCTCGAGCTGGAGACCAC	632
Q	y	201	GlnGluProSerValGluThrAlaValGlyGluSerAlaSerLysGluSerGluLeuLys	220
D	b	633	CAGGAGCCCAGTGTGGAGACTGCCGTGGAGAGTCAGCATCCAAGAAATGAGCTGAAG	692
Q	y	221	GlnSerThrGluLysGlnGluGlyThrLeuLysGlnGluGlnSerSerThrGluIlePro	240
D	b	693	CAATCCACAGAGAACAGAGGCACCTTGAAACAAGACAGAGCAGCACAGAAATCCCC	752
Q	y	241	LeuGlnAlaGluSerAspGlnAlaAlaGluGluGluAlaLysAspGluGlyGluLys	260
D	b	753	CTTCAAGCCGAATCTCATCAAGCGGTGAGGAAGAAGCCAAAGATGAAGAGAGAATAA	812
Q	y	261	GlnGluLysGluProThrlYsSerProGluSerProSerSerProValAnsSerGluThr	280
D	b	813	CAAGAGAAAGAGCCCAACCAAGTCCCCAGAAATCCCCGAGCAGCCCACTCAA	872
Q	y	281	ThrSerSerPheLysLysPheThrHisGlyTrpAlaGlyTrpArgLysLysThrSer	300
D	b	873	ACATCTCTCTTCAAGAAGTTCTTCACTCAGGTGTGGCCGGCTGGCCCAAGAAGACCAG	932
Q	y	301	PheLysLysSerLysGluAspLeuGluThrAlaGluLysArgLysGluGlnGluAla	320
D	b	933	TTCAGAANAATCAAAAGAGGATGATCTGGAAACTGCCGAGAAGAAAGGAGCAAGAGCA	992
Q	y	321	GluLysValAspGluGluLysGluLysThrGluProAlaSerGluGluGlnGluPro	340
D	b	993	GA AAAAGTAGACAGGAGAAAGAAAGAAACACAGAGCAGCCTCGAGGAGCAGGAGCCG	1052
Q	y	341	AlaGluAspThrAspGlnAlaArgLeuSerAlaAspTyrGluLysValGluLeuProLeu	360
D	b	1053	GCAGAAGACACAGCCAGGCCAGGTGTGACGACAGCTTACAGAAGGTGGAGCTGCCTTTG	1112
Q	y	361	GluAspGlnValGlyAspLeuGluAlaSerSerGluGluLysCysAlaProLeuAlaThr	380
D	b	1113	GAAGACCAGGTTGGTGACCTCGAGGCGCATCGTCAGAGGAGAAAGTGTGCTCTCTTTGGCAACG	1172
Q	y	381	GluValPheAspGluLysMetGluAlaHisGlnGluValValAlaGluValHisValSer	400
D	b	1173	GAAGTGTGTGATGAGAAGATGGAAGCCCAAGAAAGTTCTTCAGAGGTCACAGTCAGC	1232
Q	y	401	ThrValGluLysThrGluGluGlnGlnGlyGlyGlyGluAlaGluGlyGlyValVal	420
D	b	1233	ACCGTGGAGAAGACAGAGGAGGACGAGGAGGAGGAGGAGGAGGCTGAAGGGGGCGTGGT	1292
Q	y	421	ValGluGlyThrGlyGluSerLeuProProGluLysLeuAlaGluProGlnGluValPro	440
D	b	1293	GTAGAAGGAACAGGAGAAATCTTGGCCCTGAGAAACTGGCTGAGCCCCCAGGAGGTCCCC	1352
Q	y	441	GlnGluAlaGluProAlaGluGluLeuMetLysSerArgGluMetCysValSerGlyGly	460
D	b	1353	CAGGAAGCTGAGCCTGCTGAGGAGCTGATGAACAGCAGACAGATGTGTCTCTCTGGAGGA	1412
Q	y	461	AspHisThrGlnLeuThrAspLeuSerProGluGluLysThrLeuProLysHisProGlu	480
D	b	1413	GACCACTCAACTGACAGACCTTAAGTCTCTGAAGAGAAGACGCTGCCCAACACCCAGAA	1472
Q	y	481	GlyIleValSerGluValGluMetLeuSerSerGlnGluArgIleLysValGlnGlySer	500
D	b	1473	GGCATTTGTCAGTGAGGTGGAGATGCTCTCTCTCAGAAAGAAATCAAGGTACAGGGAAGT	1532
Q	y	501	ProLeuLysLysLeuPheSerSerSerGlyLeuLysLysLeuSerGlyLysLysGlnLys	520
D	b	1533	CCCTTGAAGAACTCTTCAGTAGCTCAGGCTTAAGAGAGCTGTCTGGAGAAGACAGAG	1592
Q	y	521	GlyLysArgGlyGlyGlyAspGluGluProGlyGluTyrGlnHisIleHisThrGlu	540
D	b	1593	GGAAACAGGAGGCTGGGGGAGACGAAGAGCCTGGGAATATCAACACATTCACACCCGAA	1652

QY	541	SerrProgluSerAlaaspGluGlnlysglyGluSerSerAlaSerSerProgluGluPro	560
Db	1653	TCCCCAGAGAGTGTGTATGAGCAAGAGGGGAGAGAGCTCTGCTCGTCCCGAGAGAGCTT	1712
QY	561	GluGluThrThrCysLeuGlnlysglyProLeuGlnAlaProGlnAspGlyGluAlaGlu	580
Db	1713	GAGGAGACCAAGTGTCTGGAGAAAGGGCCGCTGGAGCACCCCGAGATGGGGAACTGAG	1772
QY	581	GluGlyThrThrSerAspGlyGlnlyslsArgGluGlyIleThrProTrpAlaSerPhe	600
Db	1773	GAAAGAACTACTTCCGATGGAGAGAAAGAGAGGATCACTCCCTGGGCACTTTC	1832
QY	601	LysLysMetValThrProLysLysArgValArgArgProSerGluSerAspLysGluGlu	620
Db	1833	AAAAAGATGGTGAAACCCCAAGAAACGGGCTCCGAAAGCTTTCGAGAGTGACAAAGAGGAA	1892
QY	621	GluLeuGlnLysValLysSerAlaThrLeuSerSerThrAspSerThrValSerGluMet	640
Db	1893	GAGCTGGAGAAAGTCCAAGAGGGCCACTTGTCTCTCCACTGATAGCAAGTGTCAAGAAATG	1952
QY	641	GlnAspGluValLysThrValGlyGluGluGlnLysProGluGluProLysArgArgVal	660
Db	1953	CAAGTCAAGTCAAAACTGTTGGTGAGGAACAABAACCAGAGGAACCAAGCTGAGGGTG	2012
QY	661	AspThrSerValSerTrpGluAlaLeuIleCysValGlySerSerLysLysArgAlaArg	680
Db	2013	GATACCTCAGTGTCTTGGGGAAGCACTGATTTGGTGGGATCTATCCAAAGAGAGCAAG	2072
QY	681	LysAlaSerSerSerAspAspGluGlyGlyProArgThrLeuGlyGlyAspSerHisArg	700
Db	2073	AAGGATCTCTTCAATGATGAAAGAGGGGCCAAGAGCACTGGAGGGAGCACTCAAGA	2132
QY	701	AlaGluGlnAlaSerLysAspLysGluAlaGlyThrAspAlaValProAlaSerThrGln	720
Db	2133	GCAAGAGAGGCCACACAAGACAAGAAAGCCGGAACAAGACGCTGTTCTCCAGCACCCAG	2192
QY	721	GluGlnAspGlnAlaGlnGlySerSerSerProGluProAlaGlySerProSerGluGly	740
Db	2193	GAGCAGGACCAAGCGCAAGGAAGTTCTCTCCAGCACAGCGGGAAGCCCTTCCGAAGGG	2252
QY	741	GluGlyValSerThrTrpGluSerPheLysArgLeuValThrProArgLysLysSerLys	760
Db	2253	GAAAGTGTCTCACTTGGGAGTCAATTTAAAGATTTGATCACTCCAGAAAAAAAACCAAG	2312
QY	761	SerLysLeuGlnGlnLysAlaGluAspSerSerValGluGlnLeuSerThrGluIleGlu	780
Db	2313	TCAAACTCGAAGAGAAAGCCGAAGACTCTAGTGTAGAGCACTTGTCCACTCACTAGTCAAA	2372
QY	781	ProSerArgGluGlnSerTrpValSerIleLysLysPheIleProGlyValArgArgLys	800
Db	2373	CCGAGTAGAGAAAGATCTTGGGTTTCCATTAAAGAAATTCATCCCGGACGGCGGAAGAA	2432
QY	801	ArgAlaAspGlyLysGlnGlnGlnAlaThrValGluAspSerGlyProValGluIleAsn	820
Db	2433	AGGGCAGACGGGAAGCAAGAAACAAGCACTGTGAAGACTCAAGGGCCAACTGGAGATTAAT	2492
QY	821	GluAspAspProAsnValProAlaValValProLeuSerGlyThrAsnAlaValGluArg	840
Db	2493	GAGGACCACTTAATGTCTCCAGCCGTGTCTCTGTCTGTCAATTAATCACTGAGAGGG	2552
QY	841	GluLysMetGluAlaGlnGlyAsnThrGluLeuProGlnLeuLeuGlyAlaValTyrrVal	860
Db	2553	GAGAAAGATGAAGACCCAGGGGGAATAGCGAGCTCCCAAGCTGTGGGGGGCTGTGAAGTG	2612
QY	861	SerGluGluLeuSerLysThrLeuValHisThrValSerValAlaValIleAspGlyThr	880
Db	2613	TCCGAGGAGCTCAATTAAGACTCTGTGTCCACACTGTAGATGTGCGAGCTCAATGATGGGACC	2672
QY	881	ArgAlaValThrSerValGluGluLysArgSerProSerTrpIleSerAlaSerValThrGlu	900
Db	2673	AGGGCAGTCAACATGTGTCAAGAACCGGTCTCTCTTGATGATATACCGCTTCGGTAAACAAG	2732

QY	901	ProeugluThrsThraLaaglValametProProValIgluIgluValThrgluIysasp	920
Db	2733	CCTCTTAAACAACACGGGAGAGCCACCTGTTGAAGAGCTCACGAAAAAGAC	2792
QY	921	IleIleLaaglIugluThrProValLeuThrgluThrLeuProIugluIlyAspAlHis	940
Db	2793	ATCATTCAGAAAGAACTCTGTGTCTCACCCAGAGTTTACAGAGGTTAAAGATGCCAT	2852
QY	941	AspAspMetValThrsSerIgluValAspPheThrsSerIgluValaValThraIarThrgluThr	960
Db	2853	GACGACATGGGTCAACAGTGAAGGTGATTCCCTCAGAAAGCTGTGACAGCCACAGAGACC	2912
QY	961	SerIgluValaLeuThrgluThrgluValaThrgluValaSerIgluValaIgluIugluThrThrasp	980
Db	2913	TCAAGAGCTCTCCGTACTGAAGAGTTACCGAAGATCGGGGGCCGGAAGAACCAACAGAC	2972
QY	981	MetValSerAlaValSerIgluLeuThraspSerProAspThrThrgluValaIarThraPro	1000
Db	2973	ATGGTGTCCCGCAGTTTCCAGCTGACTGCTCCAGACACCAAGAGAAAGCCACCCCA	3032
QY	1001	ValIgluIgluValIgluSerIgluValLeuAspThrgluIgluIgluValargIarThrgluAla	1020
Db	3033	GTTCCAGGAGTGAAGAGTGGTGTGCTAGATACAGAAAGAGAGAGCCGACAGCCAGGCC	3092
QY	1021	IleLeuIgluAlaValaAlaAspIysValIysIgluIgluSerIgluValProaIarThrgluThr	1040
Db	3093	ATTCCTCAAGCCGTGTGACGACAAAGTGAAGAGAGGTCCAGGTGCCGCAACCCAGACT	3152
QY	1041	ValIgluThrgluThrgluSerIysAlaLeuIgluIysValaIgluIgluValIgluIgluIgluAspSerIglu	1060
Db	3153	GTCAGAGAAACGGGGTCTCAAGCATCTGGAGAAAGTTGAGAGGTTAGAGAGAGACTCCGAA	3212
QY	1061	ValLeuAlaSerIgluIysIgluIysAspValIleCProIysIgluProValIgluIgluAlaIglu	1080
Db	3213	GTCGTGGCTTCGGAGAAAGAGAGAGCGTTATGCCAAAGAGCCGTGCCAGAGAGCTGGA	3272
QY	1081	AlaIgluIleuAlaIgluIgluSerIgluThrgluIgluAlaIarThraProIgluSerLeuIgluVal	1100
Db	3273	GCTGAGCATCTTGCACAGGGCTCTGTGAGCTGACAGGGCTACTCCAGAGAGCTTTGAAGATT	3332
QY	1101	ProIgluValThraIarAspValaAspHisValaIarThraCysIgluValaIleIysLeuIgluIglu	1120
Db	3333	CCTAAGTCAACGGAGATGAGACACTGTGCGACAGTGCACAGTTATCAAGCTCCACAG	3392
QY	1121	LeuMetIgluIgluAlaValaAlaProIgluSerSerIgluThrLeuThraspSerIgluThraAsn	1140
Db	3393	CTGATGGAACAGCGCCGTGGCCCTGTAGTCATCCGAACCTTGACACAGACAGTGAACAAT	3452
QY	1141	GlySerThraProeulAlaAspSerAspThraIaAspGlyThrgluIgluIgluIgluIgluThrIle	1160
Db	3453	GGAAGCACTCCCTTAGACAGATTCAGACATCGCAGATGGGACACAGCAAGATGAACAACATT	3512
QY	1161	AspSerIgluAspSerIysAlaThraAlaValaIargIgluIgluIgluIgluValThrgluIgluIglu	1180
Db	3513	GACGACGACAGTGAAGCACTGACGTGTGACAGAGTCAAGTCAAGAAAGAGAG	3572
QY	1181	AlaAlaThraIgluIgluIgluIgluIgluIgluIgluIgluIgluIgluIgluIgluIgluIgluIglu	1200
Db	3573	GGCGCTACTGTCAGAAAGAGAGCTTGACACATCACTTAATATGTTCCAGCCAGGAA	3632
QY	1201	GluHisIgluIgluIgluIgluIgluIgluIgluIgluIgluIgluIgluIgluIgluIgluIgluIglu	1220
Db	3633	GAACTATGGGGAAGAACAGAGAGATGTTCTTGAACCTACACAGACAGACTTACTCTCT	3692
QY	1221	AlaAlaValaProValLeuAlaIlyThrgluValaIgluIgluIgluIgluIgluIgluIgluIgluIglu	1240
Db	3693	GCAAGCCGTGCCCTTCTGGCAAGACTGTGAGGTGGGTCAAGAGGTGAAGTTGACTGTGGTTG	3752
QY	1241	AspGlyIgluIysValIysIgluIgluIgluIgluIgluIgluIgluIgluIgluIgluIgluIgluIglu	1260
Db	3753	GATGGAGAAAAGTCAAGAGAAAGACAGAGAGGTGTTGTACATCTGTGACCCAAACGTGAA	3812
QY	1261	LysAlaIlaAspValThrTyraAspSerIgluValaMetGlyValaIlaIgluIgluIgluIgluIglu	1280

```

Db 3813  AGGCTCTGATGACATATGACAGTGAAGTATGATGGAGTGGCCGGTGTCCAGAAAG 3872
Qy 1281  GluSerThrGluValGlnSerLeuSerLeuGluGluGlyGluMetGluThrAspValGlu 1300
Db 3873  GAGAGTACTGAGTGCAGAGCTTAGCTGGAGGAGGAGATGGAATGACGTTGAA 3932
Qy 1301  LysGluLysArgGluThrLysProGluGlnValSerGluGluGluGlnGluThrAla 1320
Db 3933  AAGGAGAAAGGGAGACAAAGCCAGACAGCAAGTGTAGTGAAGAGGTGACGAGAAACAGCC 3992
Qy 1321  AlaProGluHisGluGlyThrTyrgLysProValLeuThrLeuAspMetProSerSer 1340
Db 3993  GCTCTGAGCATGAGAGAACCTACGGAGCCAGTCTCTGACACTTGTACATGCCAGCTCA 4052
Qy 1341  GluArgGlyLysAlaLeuGlySerLeuGlySerProSerLeuProAspGlnAspLys 1360
Db 4053  GAGAGGGGAAGGACACTGGAGAGCTTGGAGGAGCCCTTCTCTCCAGACCAAGCAAA 4112
Qy 1361  AlaGlyCysIleGluValGlnValGlnSerLeuAspThrThrValThrGlnThrAlaGlu 1380
Db 4113  GCAGGTTGCATAGAGTTCAAGTTCAAGGCTGGACACACAGTCACTCAACAGCAGAA 4172
Qy 1381  AlaValGluLysValIleGluThrValValIleSerGluThrGlyGluSerProGluCys 1400
Db 4173  GCTGTGAAAGAGGTATAGAAACGGTTGTGATTTTCAGAGACAGGTGAAGTCCAGAGTGT 4232
Qy 1401  ValGlyAlaHisLeuLeuProAlaGluLysSerSerAlaThrGlyGlyHisTrrThrLeu 1420
Db 4233  GTAGGTGCACACTATTACAGCTGAGAGTCTCTCCACGGGTGGCCACTGACTCTT 4292
Qy 1421  GlnHisAlaGluAspThrValProLeuGlyProGluSerGlnAlaGluSerIleProile 1440
Db 4293  CAGCATGCAGAGGACACGGTACCCCTGGGCGCTGAGTCTCAGGCAGAAATCCATCCCAATC 4352
Qy 1441  IleValThrProAlaProGluSerThrLeuHisProAspLeuGlnGlyIleSerAla 1460
Db 4353  ATAGTAACCTCTGCTCTGAAAGACCCCTATCCTCCTGACCTACAGGAGAAATAAGCGCA 4412
Qy 1461  SerGlnArgGluArgSerGluGluAspLysProAspAlaGlyProAspAlaAspGly 1480
Db 4413  TCCAGAGAGAGCGATCAGAGGAGAGACAGCCAGATGCTGTCTGTATGCTGACGGC 4472
Qy 1481  LysGluSerThrAlaIleGluLysValLeuLysAlaGluProGluIleLeuGluLeuGlu 1500
Db 4473  AAGGAGAGTACAGCAATCGAAAGATCCTCAAGGCTGAACCTGAGATCCTGGAACCTGAG 4532
Qy 1501  SerLysSerAsnLysIleValLeuAsnValIleGlnThrAlaValAspGlnPheAlaArg 1520
Db 4533  AGTAAGAGCAACAAAGATTGCTGAACCTCATTTCAGACAGCCGTTGACAGTTCCGACGT 4592
Qy 1521  ThrGluThrAlaProGluThrHisAlaTyAspSerGlnThrGlnValProAlaCysArg 1540
Db 4593  ACAGAAACACCCCCGAAACTCATGCTTATGATTACAGACCAGAGTTCTGTGATGAGG 4652
Qy 1541  LeuAspSerArgGluProAsnArgCysTrpThrLysMetLysAspAlaLysMetLysHis 1560
Db 4653  CTTGACAGCAGGAGGCCAACACAGATGCTGGACAAATGAAAGATGCCAAGATGAACAC 4712
Qy 1561  ProValProGlnProArgGluAspLeuGlnValLeuThrValLeuGluAlaTrpAlaGln 1580
Db 4713  CCAGTGGCGAGGCCAGAGAGGACTTCAAGTCTGACCGCTTCTGGAGGCGATGGCTCAG 4772
Qy 1581  ProArgLysCysLeuProArgLeuGlnLeuLysAlaProValSerLys 1596
Db 4773  CCTCGAAATGCTTGGCGCGCTTGCAGTTGAAAGCGCCGGTGTCAAAG 4820

```

RESULT 3

US-08-665-401-3

; Sequence 3, Application US/08665401

; GENERAL INFORMATION:

; APPLICANT: Gelman, Irwin H.

; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE

```

; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10112-0228
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/665,401
; FILING DATE: 18-JUN-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Richard S
; REGISTRATION NUMBER: 26,154
; REFERENCE/DOCKET NUMBER: A30558 - 165/34008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-408-2558
; TELEFAX: 212-765-2519
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6160 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; US-08-665-401-3

Alignment Scores:
Pred. No.: 0 Length: 6160
Score: 8073.00 Matches: 1596
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-902-432-4 (1-1596) x US-08-665-401-3 (1-6160)

Qy 1 MetGlyAlaGlySerSerThrGluGlnArgSerProGluGlnProAlaGlySerAspThr 20
Db 33 ATGGGCGCAGGCGAGTTCACCCGAGCAGCGAGCCCGGAGCGAGCGACG 92
Qy 21 ProSerGluLeuValLeuSerGlyHisGlyProAlaAlaGluAlaSerGlyAlaAlaGly 40
Db 93 CCGAGCGAGCTGTGTCTCAGTGGCCATGGCCGCGCAGCTGAAGCTCGGAGCGAGCTGA 152
Qy 41 AspProAlaAspAlaAspProAlaThrLysLeuProGlnLysAsnGlyGlnLeuSerSer 60
Db 153 GACCCCGCGCAGCGGACCCCGCCACCAAGCTCCACAGAGAATGGCCAGCTGTCTTCT 212
Qy 61 ValAsnGlyValAlaGluGlnGlyAspValHisValGlnGluGlnGlnGluGlyGln 80
Db 213 GTCACCGCGGTAGCTGAACAGAGAGATGTCATGTCCAGAGAGAAACCAAGGAGGGGCG 272
Qy 81 GluGluGluValValAspGluAspValGlyGlnArgGluSerGluAspValArgGluLys 100
Db 273 GAGCAAGAGAGTGTGTGATGAGGATGTTGGACAGCAGAGTCAAGAGATGTGAGAGAAAA 332
Qy 101 AspArgValGluGluMetAlaAlaAsnSerThrAlaValGluAspIleThrLysAspGly 120

```

Db	333	GACCGAGTTGAAGAAATGGCGGCCCAACTCCACAGCTGTTGAAGATATTCACAAAGATGGG	392
Qy	121	GInGInGInUThrSerGInUleIleGInGInIleProAlaSerGInUAsnAsnValGInUInU	140
Db	393	CAGAGAGAGACATCAGAAATTAATTGAACAGATCTCCGTCTTCAGAAAACAAATGTGGAAATA	452
Qy	141	MetValGInProAlaGInUSeRGInAlaAsnAspValGlyPheUlysValPheUlysPhe	160
Db	453	ATGGTACAGCTCGCTGAGTCCCGACGGCTTAAGATGTTGGCTTCAGAAAGATATTTAAATTT	512
Qy	161	ValGlyPheUlysPheMetThrValUlysValAspUlysAsnGInUlysSerPheThrValGInUleu	180
Db	513	GTTGGTTTAAATTCACGGTGGAAGAGATTAATAAGAAAGTCAGATCTGTCCACACTA	572
Qy	181	LeuThrValUlysValAspGInUglyGInUglyAlaGlnAlaSerValGlyValaGlyAspHis	200
Db	573	CTCACTGTCAAGAAAGATGAAGGGGAAGGGGCAAGAGCCCTGTGCGAGCTGGACACACAC	632
Qy	201	GInGInUProSerValGInUThrAlaValGlyUSeRAlaSerUlysGInUSeRGlUleUlys	220
Db	633	CAGAGAGCCAGTGTGGAGACTCGCGTGGAGAGTCCAGTCCAAAGAAAGTGAAGCTGAAG	692
Qy	221	GInSerThrGInUlyGInGInUglyThrLeuUlyGInGInUInSerSerThrGInUlePro	240
Db	693	CAATCCACAGAGAAACAGAAAGGACCCCTTAACCAAGAACAGAGCAGCACAAAATATCCCTC	752
Qy	241	LeuGlnAlaGInUSeRAspGInAlaAlaGInUglyAlaAlaUlysAspGInUglyGInUlys	260
Db	753	CTTCAGAGCCGAATCTGATCAAGCGGCTGAGAGAAAGGCCAAAGATGAAAGAAATAAAAA	812
Qy	261	GInGInUlySGInUProThrUlysSerProGInUSeRProSerSerProValAsnSerGInUThr	280
Db	813	CAAGAGAAAGAGCCACCAAGATCCCCAGAAATCCCCAGAGAGCCCAAGTCAACAGTAGACAA	872
Qy	281	ThrSerSerPheUlysPhePheThrHisGlyThrAlaGlyThrPArgUlySerThSer	300
Db	873	ACATTTCTTCAAGAAATTTCTACTCAAGTGTGGCGGGCTGGGCAAGAAAGCAAGC	932
Qy	301	PheUlysSerUlysGInUAspAspLeuGInUThrAlaGInUlyeHrgUySGInUInGInUAla	320
Db	933	TTCAAGAAATCAAAAGAGATGATCTTGAAATCTGCCGAAAGAAAGAGCAAGCAAGAGCA	992
Qy	321	GInUlysValaAspGInGInUglyUlysGInUlySerGInUProAlaSerGInUglyGInUInUPro	340
Db	993	GAAAAAGTAGACGAGAGAAAGAAAGAAAGAACAGAGCCAGGCTCGGAGAGCAGAGAGCCG	1052
Qy	341	AlaGlnAspThrAspGInAlaAlaGlyUSeRAlaAspUrgGInUlysValGInUleuProUen	360
Db	1053	GCAGAAAGCACAGACCAAGGCCAGGGTTGTACGAGACTACAGAAAGGTGAGAGCTGCTTTTG	1112
Qy	361	GInUAspGInUValGInUAspLeuGlnAlaSerSerGInUlyeGValaProUenAlaThr	380
Db	1113	GAAAGCCAGGTGGTGACCTGGAGGCACTCGTCAAGAGAGAAAGTGTCTCTTTGGCAACG	1172
Qy	381	GInUAlaPheAspGInUlyeMetGlnAlaHisGInGInUValaValaGlnUAlaHisValSer	400
Db	1173	GAAAGTTTGATGAGAGATGGAAGCCCAACAAAGATGTTTCAAGAGGTCCACAGTGAAGC	1232
Qy	401	ThrValGInUlySerThrGInGInUglyGInGInUglyGlyGInUAlaGInUglyGlyValaVal	420
Db	1233	ACCGTGGAAGAAACAGAGAGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1292
Qy	421	ValGInUglyThrGInUglyUSeRLeuProGInUlyeUenAlaGInUProGInGInUValaPro	440
Db	1293	GTAGAAGGACAGAGGAATCTTCTCCCTCGAGAAACTGGCTGAAGCCCAAGAGGTCCCC	1352
Qy	441	GInGInUAlaGInUProAlaGInUleuUlyeMetUlysSerArgGInUmetCysValSerGlyGly	460
Db	1353	CAGGAAGCTGAAGCCTCGTGAGAGACTGATGAAGGACGAGAAGATGTGTCTCTGAAGGA	1412
Qy	461	AspHisThrGInUleuThrAspUleuSerProGInUglyUlysThrUleuProUlyHisProGInU	480
Db	1413	GACCACTCTCACTGACAGACTTAATAGTCTGGAAGAGAGAGCGTGGCCAAACACCAAGAA	1472

QY	461	GLYLLEVALSERGIUVALGIUMLEUSERSERGIUVALGIULEYSEVALGINGLYSER	500
Db	1473	GGCATTTGTCAGTGGTGGTGGAGATGCTGCTCTCCAGCAAGAAAGATCAAGATTCAGAGGAAT	1532
QY	501	PROLEUYSILYSLEUPHESERSESERGIYLEUYSILYSEUSERGIYLYLYLEGLNLS	520
Db	1533	CCCTTGAAGAAACTCTTCAGTACTACGGCTTTAAAGAGCTGTCTGGGAAAGAACACAGAG	1592
QY	521	GLIYLSAARGIYGLIYGLIYLAEPGIUGIUPROGIYGLIUTYRGINHIEIETHRGINU	540
Db	1593	GGGAAACAGAGAGGTGGGGAGACCAAGAGCCTCGAGAAATTCACACATTCACACCGAA	1652
QY	541	SERPROGIUSERALAEAPGIUGIUNYSGIYGIUSERSERALAESESEPROGIUGIUPRO	560
Db	1653	TCCCAAGAGAGTGTGATGAGCAGAAAGGAGAGAGCTGCTGCTGCTCCCGAGGAGCTT	1712
QY	561	GIUGIUTHIRTHRCYSEUGIULYSGIYPROLEUGIUNIAEPROGINAEPGIUGIUNIAGLU	580
Db	1713	GAGAGACACACGTGTCTCGAGAAAGGGCCGCTGAGAACGCCACAGAGAGGGGAAAGCTGAG	1772
QY	581	GIUGIUTHIRTHRSERAPGIYGIULYLSLVSARVGIUGIULIETHPROTPALAESEPH	600
Db	1773	GAAAGCACTACTTCCGATGAGAGAAAGAAAGAGAGGAGATCACTCTCGAGCATCTTC	1832
QY	601	LYSLYSEMETVALTHIRPOLYSLYSARVVALARGARVPROSESGIUSERASPLYSGIUGIU	620
Db	1833	AAAAAAGATGGGACACCCCAAGAAACGGGTCGGAAGACCTTCGAGATGACACAGAGAGAA	1892
QY	621	GIUENGIULYSVALYSESERALATHIRLEUSERSETHIRAPSETHIVALSERGIUMET	640
Db	1893	GAGCTGGAGAAAGGTCAAGACCGCCACCTTGCTCTCCACGTAGACACAGTCAAGAAATG	1952
QY	641	GLINAPGIUVALYSTRTHRVALIGIYUGIUGIUNYSPROGIUNGIUNIPOLYSAIRARVVAL	660
Db	1953	CAGATGAAGTCABAACTGTTGTGTGGAAACAAAGCCAGAGAAACAAAGCGTGAAGGCTG	2012
QY	661	ASPTHIRSERVALSERTPGIUNIALEUIECYSEVALGISESERLYSEARVALAARG	680
Db	2013	GATTCCTCAGTGTCTTGGAAAGACGATTTGTGTGTGATCATCCAAAGAGAGCAAG	2072
QY	681	LYSALASERSESERAPAPPGIUGIYGIYPROARGTHIRLEUGIYGLYASPSERHISARG	700
Db	2073	AAGGATCTCTTTCAGATGATGAAAGAGGGGCAAGGACACTGGAGGGGAGCAGTCACAGA	2132
QY	701	ALAGIUGIUNIASEIRYSAEPLYSGLIYAGIYTHIRASPALAVALPROALAESETHGIN	720
Db	2133	GCAAGAGAGGCGACACAAAGCAAAAGAGCCGGAACGAGCGTGTCTCCGACGACCCAG	2192
QY	721	GIUGIUNAPGINALAGIINGIYSE	740
Db	2193	GAGCAGACCAAGGCGCAAGAACTTCTCAACCCGAGCCAGCGGGAAGCCCTTCCGAAAGG	2252
QY	741	GIUGIYALSERTHIRTPGIUSERPHELYSARGLEUVALTHIRPROARGIYLYSESELYS	760
Db	2253	GAAAGTGTCTCACTTGGAGGTCACTTTAAAGATTAGTCATCCCAAGAAAAAAATCCAG	2312
QY	761	SERLYSEUGIUGIULYVALAGIUSAPSE	780
Db	2313	TCAAAACTGGAAGAAAGCCGGAAGCTTAGTGNAGACCAATTTCCACTGAGATGCCAA	2372
QY	781	PROSEARVGIUGIUSERTRPVALSERILEYLYSPHEPIIEPRIGIYAIRARGIYLYS	800
Db	2373	CCGAGTGAAGAAAGATCTTGGGTTTCATTTAAGAAATTCATTCGCCGAGCCGGCGAAGAA	2432
QY	801	ARGIYLAEPGIYLYSGIUGIUNIALATRVALGIUSAPSE	820
Db	2433	AGGGCAGCCGGGAAGCAAGAACCAAGCCTGTGGAAAGACTCAAGGGCCAGTGGAGATAAT	2492
QY	821	GLIUSAPPROABENVALPROALAVALVALPROLEUSERGIUTYTRASNALAVAGIUNARG	840
Db	2493	GAGGACGACCTTAATGTCCAGCGCTGTGGCTCTGTCTGAATTAAGCAGTGGAGAGG	2552

QY 841 GluLysMetGluAlaGlnGlyAsnThrGluLeuProGlnLeuLeuGlyAlaValTyrVal 860
DB 2553 GAGAAGATGGAAGCCCAAGGGGAATACAGGAGTGCCTCCAGCTGTGGGGGCTGTATCGTG 2612
QY 861 SerGluGluLeuSerLysThrLeuValHisThrValSerValAlaValIleAspGlyThr 880
DB 2613 TCCGAGAGCTCAGTAAGACTCTGGTCCACACTGTGAGTGTCCGACGTCATTGTATGGACC 2672
QY 881 ArgAlaValThrSerValGluGluArgSerProSerTrpIleSerAlaSerValThrGlu 900
DB 2673 AGGCGAGTCACCAAGTTCGAAGAGCGGTCTCTTCGTGGATATCCCGTCTCCGTAACAGAA 2732
QY 901 ProLeuGluHisThrAlaGlyGluAlaMetProProValGluGluValThrGluLysAsp 920
DB 2733 CCTCTTGAACACACAGCGGAGAGCCATGCCACCTCTTGAAGAGGTCACTGAAAAAGAC 2792
QY 921 IleIleAlaGluThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHis 940
DB 2793 ATCATTCAGAGNAACCTCTGTGCTCACCCAGACGTTACAGAGGGTAAAGATGCCCAT 2852
QY 941 AspAspMetValThrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThr 960
DB 2853 GACGACATGCTCACCACTGAAGTGAATTCACCTCAGAAGCTGTGACAGCCACAGAGACC 2912
QY 961 SerGluAlaLeuArgThrGluGluValThrGluAlaSerGlyAlaGluGluThrThrAsp 980
DB 2913 TCAGAGGCTCTCCGTACTGAAGAAGTTTACCAGAACATTCGGGGGCCGAAGACCCACAGAC 2972
QY 981 MetValSerAlaValSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrPro 1000
DB 2973 ATGGTGTCCGAGTTTCCCAGCTGACTGACTTCCCAGACACCACAGAGGAAGCCACCCCA 3032
QY 1001 ValGlnGluValGluSerGlyValLeuAspThrGluGluGluArgGlnThrGlnAla 1020
DB 3033 GTTCAGGAGTAGAGAGTGGTGTGCTAGATACAGAAGAGAGGCGCCAGAGCC 3092
QY 1021 IleLeuGlnAlaValAlaAspLysValLysGluGluSerGlnValProAlaThrGlnThr 1040
DB 3093 ATCCTCCAGCCGTTCAGACAAGGTGAAGAGAGTCCCAAGAGTCCCGAGTGCCTGCAACCCAGACT 3152
QY 1041 ValGlnArgThrGlySerLysAlaLeuGluLysValGluGluValGluGluAspSerGlu 1060
DB 3153 GTGCAGAGAACGGGTCAAAAGCACCTGGAGAAGTTTGGAGAGGTAGAGGAGACTCCGAA 3212
QY 1061 ValLeuAlaSerGluLysGluLysAspValMetProLysGlyProValGlnGluAlaGly 1080
DB 3213 GTGTGGCTTCGGAGAAAGAGAGGAGCTTATGCCGAAAGACCCCGTCAGGAAGCTGGA 3272
QY 1081 AlaGluHisLeuAlaGlnGlySerGluThrGlyGlnAlaThrProGluSerLeuGluVal 1100
DB 3273 GCTGAGCATCTTCACAGGGCTCTGAGACTGGACAGGCTACTCCAGAGAGCCTTGAAGTT 3332
QY 1101 ProGluValThrAlaAspValAspHisValAlaThrCysGlnValIleLysLeuGlnGln 1120
DB 3333 CCTGAAGTACGGCAGATGAGACCATGTGCCACGTCGCCAGGTATCAAGCTCCAGCAG 3392
QY 1121 LeuMetGluGlnAlaValAlaProGluSerSerGluThrLeuThrAspSerGluThrAsn 1140
DB 3393 CTGATGGAACAGGCGGTGGCCCTCGTACTCCGAAACCTTGACAGACACTGAGACAAAT 3452
QY 1141 GlySerThrProLeuAlaAspSerAspThrAlaAspGlyThrGlnGlnAspGluThrIle 1160
DB 3453 GGAAGCACCTCCCTTAGCAGATTACAGACACTGCAGATGGGACACAGCAAGATGAACCAAT 3512
QY 1161 AspSerGlnAspSerLysAlaThrAlaAlaValArgGlnSerGlnValThrGluGluGlu 1180
DB 3513 GACAGCCAGACAGATGAAGCCACTGCAGCTGTTCAGGAGTTCACAGGTTCAGAGAAGAG 3572
QY 1181 AlaAlaThrAlaGlnLysGluGluProSerThrLeuProAsnAsnValProAlaGlnGlu 1200
DB 3573 CGGGCTACTGCTCAGAAAGAGGAGCCCTTCACACTACTCTAATATGTTCCAGCCCAAGAA 3632
QY 1201 GluHisGlyGluGluProGlyArgAspValLeuGluProThrGlnGlnLeuThrAla 1220

DB 3633 GAACATGGGGAAAGAACCCAGGAAGAGATGTTCTTGAAACCTACACAGCAAGAGCTTACTGCT 3692
QY 1221 AlaAlaValProValLeuAlaLysThrGluValGlyGlnGluGlyGluValAspTrpLeu 1240
DB 3693 GCAGCCGTGCCCGCTTCTGGCAAGACTGAGGTGGGTCAAGAGGTGAGGTGTGACTGGTTG 3752
QY 1241 AspGlyGluLysValLysGluGlnGluValPheValHisSerGlyProAsnSerGln 1260
DB 3753 GATGGAGAAAAAGTCAAAAGAAACAGGAGGTGTTTGTACACTCTCGACCCCAACAGTCAA 3812
QY 1261 LysAlaAlaAspValThrTyrAspSerGluValMetGlyValAlaGlyCysGlnGluLys 1280
DB 3813 AAGGCTGCTGATGTGACATATGACAGTGAAGTGAATGGAGTGGCCGGGTGTTCAGGAAAAAG 3872
QY 1281 GluSerThrGluValGlnSerLeuSerLeuGluGluGluMetGluThrAspValGlu 1300
DB 3873 GAGAGTACTGAAGTGCAGAGTCTTAGCCTTGGAGAGGAGATGGAACCTGACGTTGAA 3932
QY 1301 LysGluLysArgGluThrLysProGluGlnValSerGluGlyGluGlnGluThrAla 1320
DB 3933 AAGCAGAAAAGGGAGACAAAGCCAGAGCAAGTGAAGTGAAGAGGTGAGCAGGAAACAGCC 3992
QY 1321 AlaProGluHisGluGlyThrTyrGlyLysProValLeuThrLeuAspMetProSerSer 1340
DB 3993 GCTCTCTGAGCATGAAGGAAACCTACGGGAAGCCAGTCTTGACACTTGACATGCCAGTCA 4052
QY 1341 GluArgGlyLysAlaLeuGlySerLeuGlyGlySerProSerLeuProAspGlnAspLys 1360
DB 4053 GAGAGGGGAAGGACACTGGGAAGCCTTGGAGGAAGCCCTTCTCCAGACCAAGACAAA 4112
QY 1361 AlaGlyCysIleGluValGlnValGlnSerLeuAspThrThrValThrGlnThrAlaGlu 1380
DB 4113 GCAGGTTGCATAGAGTTCAGGTTCAAGCTTCAAGCTGGACACAAACAGTCACTCAAAACAGCAGAA 4172
QY 1381 AlaValGluLysValIleGluThrValValIleSerGluThrGlyGluSerProGluCys 1400
DB 4173 GCTGTGCAAAAAGGTCATAGAAAACGGTGTGATTTTCAGAGACAGGTGAAAGTCCAGAGTGT 4232
QY 1401 ValGlyValHisLeuLeuProAlaGluLysSerSerAlaThrGlyGlyHisTrpThrLeu 1420
DB 4233 GTAGGTGACACACTTATACAGCTGAGAAGTCTCTGCAACGGGTGGCCACTGGACTCTT 4292
QY 1421 GlnHisAlaGluAspThrValProLeuGlyProGluSerGlnAlaGluSerIleProIle 1440
DB 4293 CAGCATGCAGAGGACACAGCTACCCCTGGGGCCTGAGTCTCAGGCAGACATCCATCCCAATC 4352
QY 1441 IleValThrProAlaProGluSerThrLeuHisProAspLeuGlnGlyGluIleSerAla 1460
DB 4353 ATAGTAACTCTCTCTCTGAAAACACCTTACATCTCTGACCTACAGGAGAAATTAAGCGCA 4412
QY 1461 SerGlnArgGluArgSerGluGluAspLysProAspAlaGlyProAspAlaAspGly 1480
DB 4413 TCCAGAGAGAGGGATCAGAGGAAGAGGACAGCCAGATGCTGGTCTCTGATGCTGACGGC 4472
QY 1481 LysGluSerThrAlaIleGluLysValLeuLysAlaGluProGluIleLeuGluLeuGlu 1500
DB 4473 AAGGAGGTACAGCAATCGAAAAAGTCTCAAGGCTGAACCTGAGATCTCTGGAACCTTGAG 4532
QY 1501 SerLysSerAsnLysIleValLeuAsnValIleGlnThrAlaValAspGlnPheAlaArg 1520
DB 4533 AGTAAGAGCAACAAGATTGTGCTGAACGTCATTTCAGACAGCCGTTTACCAAGTTCGCACGT 4592
QY 1521 ThrGluThrAlaProGluThrHisAlaTyrAspSerGlnThrGlnValProAlaCysArg 1540
DB 4593 ACAGAAACAGCCCCGAAACTCATGCTTATGATTTCACAGACCCAGGTCTCTGATCATCAG 4652
QY 1541 LeuAspSerArgGluProAsnArgCysTrpThrLysMetLysAspAlaLysMetLysHis 1560
DB 4653 CTTGACAGCAGGAGGCCCAACAGATGCTGGACAAAATGAAGATGCCAAGATGAAACAC 4712
QY 1561 ProValProGlnProArgGluAspLeuGlnValLeuThrValLeuGluAlaTrpAlaGln 1580

Db 4773 CCAAGTCCGACGACGAGAGAGACTTGCAAGCTTGACCCGTTCTGGAGGCATGGGCTCAG 4772
Qy 1581 ProArqlyscysleuProArqleuInleuylsaIProvalSerlys 1596
Db 4773 CCTCGGAATGCTTGCCTGCGCTTGACGATTGAAAGCGCCGGTGTCAAG 4820
RESULT 4
US-09-902-432-3
; Sequence 3, Application US/09902432
; GENERAL INFORMATION:
; APPLICANT: Irwin H. Gelman
; APPLICANT: Susan G. Jaken
; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
; FILE REFERENCE: A30558-A-FWC-A 070156.0597
; CURRENT FILING DATE: 2002-04-08
; PRIOR FILING DATE: 1997-11-25
; PRIOR FILING DATE: 1997-11-25
; PRIOR APPLICATION NUMBER: 08/665,401
; PRIOR FILING DATE: 1996-06-18
; PRIOR APPLICATION NUMBER: 08/635,121
; PRIOR FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 6160
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-09-902-432-3
Alignment Scores:
Pred. No.: 0 Length: 6160
Score: 8073.00 Matches: 1596
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
US-09-902-432-4 (1-1596) x US-09-902-432-3 (1-6160)
Qy 1 MetGlyAlaGlySerSerThrgluGlnArgSerProgluGlnProAlaGlySerAepThr 20
Db 33 ATGGGCGCGAGCATTCACCGACGACGCGAGCCCGGAGCAGCCGCGGAGCGACAG 92
Qy 21 ProSerGluLeuValLeuSerGlyHisGlyProAlaAlaGluAlaSerGlyAlaAlaGly 40
Db 93 CGAGGAGACTGGTGTCTCAGTGGCCATGGGCCCGGACGCTGAAACCTCTGGGAGCAGCTGGA 152
Qy 41 AspProAlaAspAlaAspProAlaThrIlySleuProGluIlyAsnGlyIlyInleuSerSer 60
Db 153 GACCCCGCGAGCGGACCCCGCCACCAAGCTCCACAGAAAGATGCGCAGCTGTCTTCT 212
Qy 61 ValAsnGlyValAlaGluGlnGlyAspValHisValGlnGluIlyAsnGlnGluGlyIn 80
Db 213 GTCAACGGCTGACTTAACAAGAGATGTCTCATGTCCAAAGAGAAAAACAGAGGGGCGAG 272
Qy 81 GluGluGluValValAspGluAspValGlyGlnArgGluSerGluAspValArgGlyIly 100
Db 273 GAGGAAGAAAGTCTTATGATGAGATGTTGGACAGAGAGTCAAGATGTGAGAAAAA 332
Qy 101 AspArgValAlaGluIlyMetAlaAlaAsnSerThrAlaValGluAspIlyThrIlyAspGly 120
Db 333 GACCGAGTTGAAGAAATGGCGGCCCACTCCACAGCTGTGAAATATCACAAAAGGATGGG 392
Qy 121 GlnGluGluThrSerGluIlyIleGluGlnIlyProAlaSerGluAsnAsnValGluGlu 140
Db 393 CAGGAGAGAGCATCAAGAAATATTGAACAGATCCCTGCTTCAAGAAACATGTGGAGAA 452
Qy 141 MetValGlnProAlaGluSerGlnAlaAsnAspValGlyPheIlySlyValPheIlySlyPhe 160
Db 453 ATGGTACAGCGCTCTGAGTCCCAAGGCTAATGATGTTGGCTTCAAGAAAGATTTAAATTT 512
Qy 161 ValGlyPheIlySlyPheThrValIlySlyAspIlyAsnGluIlySerSerAepThrValGlnLeu 180

Db 513 GTTGGTTTAAATTCACGCTGAAGAGATATAAAATGATGATCTGTCCACTA 572
Qy 181 LeuThrValIlySlyAspGluGlyGluGlyAlaGluAlaSerValGlyAlaGlyAspHis 200
Db 573 CTCACTGTCAAGAAAGATGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 632
Qy 201 GlnGluProSerValGluThrAlaValGlyIlySerIlyAsnGlySerGluIlyLeuIlyS 220
Db 633 CAGGAGCCAGGTGGAGACTGCGGTGGAGAGTCAAGATCAAGAAAGATGAGCTGAAG 692
Qy 221 GlnSerThrGluIlySglnGluGlyThrLeuIlySglnGlnGlnSerSerThrGluIlyPro 240
Db 693 CATCCACAGAGAGAGCAAGAGAGCACCCTGAAGACAGAGAGAGAGAGAGAGAGAGAGAGAG 752
Qy 241 LeuGlnAlaGluSerAspGlnAlaAlaGluGluGluAlaIlyAspGluGlyGluIlyS 260
Db 753 CTTCAAGCCGAATCTGATCAAGCCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 812
Qy 261 GlnGluIlySglnProThrIlySerProGluSerProSerSerProValAsnSerGluThr 280
Db 813 CAAGAGAAAGAGCCCAAGTCCCAAGATCCCGAGACGCCAGTCAACAGATGAGACA 872
Qy 281 ThrSerSerPheIlySlyPhePheThrHisGlyIlyProAlaGlyIlyProAlaGlyIlySer 300
Db 873 ACATCTTCTTCAAGAGATTTCTTCACTCAAGCTTGGCCGCTGGCCGAGAGAGAGAGAGAG 932
Qy 301 PheIlySlySerIlySglnAspAspLeuGluThrAlaGluIlyArgIlyGluGlnGluAla 320
Db 933 TTCAAGAAATCAAAAGAGATGATCTGAAACTGCCGAGAGAGAGAGAGAGAGAGAGAGAGAG 992
Qy 321 GlnIlySlyValAspGluGluGluIlySglnIlySglnIlyProAlaSerGluGlnGluPro 340
Db 993 GAAAAAGTGAAG 1052
Qy 341 AlaGluAspThrAspGlnAlaArgLeuSerAlaAspIlyGlyIlyValGluIlyLeuProLeu 360
Db 1053 GCGAAGACACAGACACAGGACGAGTGTCTCAGCAGACTACAGAGAGTGAAGTGCCTTTG 1112
Qy 361 GluAspGlnValGlyAspLeuGluAlaSerSerGluIlySlyCysAlaProLeuAlaThr 380
Db 1113 GAAGACAGAGTGTGATCTGAGGAGCATGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1172
Qy 381 GluValPheAspGluIlyMetGluAlaHisGlnGluValAlaGluValHisValSer 400
Db 1173 GAAGGTGTGATGAAGAGATGAGAGCCCAAGAGAGTGTGAGAGAGTCAAGCTGAGAGC 1232
Qy 401 ThrValGluIlySglnGluGluGlnGlyIlyGlyIlyGluAlaGluGlyValValVal 420
Db 1233 ACCGTGGAGAGACAG 1292
Qy 421 ValGluGlyThrGlyGluSerLeuProGluIlySleuAlaGluProGlnGluValPro 440
Db 1293 GTAGAAGGAAACAGAGAAATCTTGGCCCTGAGAAACTGGCTGAGAGAGAGAGAGAGAG 1352
Qy 441 GlnGluAlaGluProAlaGluGluIlyMetIlySserArgGluMetCysValSerGlyGly 460
Db 1353 CAGGAAGCTGAGCTGTGAGAGAGCTGATGAAGACAGAGAGATGTGTCTCTGAGAGA 1412
Qy 461 AspHisThrGlnLeuThrAspLeuSerProgluGluIlySlyThrLeuProIlyHisProglu 480
Db 1413 GACCACTCACTCAAGACACTTAAGTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1472
Qy 481 GlyIlyValSerGluValGluMetLeuSerSerGlnGluArgIlyLeuValGlnGlySer 500
Db 1473 GGCATTGTGAGTGAAGTGAAGTGTCTCTCAAGAAAGATCAAGTCAAGGAGAGAT 1532
Qy 501 ProLeuIlySlyLeuPheSerSerSerGlyIlyLeuIlySlySglnGlySglnIlyS 520
Db 1533 CCTTGAAGAAACTCTTCAAGTGTGAGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1592
Qy 521 GlyIlyArgGlyGlyGlyValAspGluGluProGluIlyGluIlyGlnHisIlyHisThrGlu 540

Db	1593	GGGAACGAGGAGGTGGGGGAGACGAAGAGCCTGGGAGAAATACCAACACATTCACACCGAA	1655
Qy	541	SerProGluSerAlaAspGluGlnLysGlyGluSerSerAlaSerSerProGluGluPro	560
Db	1653	TCCCCAGAGNGTCTGATGAGCAGAGGGAGAGAGCTCTGCGTCTGCTCCCGAGGAGCCT	1712
Qy	561	GluGluThrThrCysLeuGluLysGlyProLeuGluAlaProGlnAspGlyGluAlaGlu	580
Db	1713	GAGGAGACCACGTGTCTGGAGAAAGGCGCTGGAAGCACCCAGGATGGGAAGCTGAG	1772
Qy	581	GluGlyThrThrSerAspGlyGluLysLysArgGluGlyIleThrProTrpAlaSerPhe	600
Db	1773	GAAGGAACCTACTCTCCATGGAGAGAGAGAGAGAGGATCATCTCCCTGGGCATCTCTC	1832
Qy	601	LysLysMetValThrProLysLysArgValArgProSerGluSerAspLysGluGlu	620
Db	1833	AAANAGATGGTGACACCCCAAGAAACGGGTCCGAAGACCTTCTGAGAGTGACAAGGAGAA	1892
Qy	621	GluLeuGluLysValLysSerAlaThrLeuSerSerThrAspSerThrValSerGluMet	640
Db	1893	GAGCTGGAGAAAGGTCAAGAGCGCCACTTGTCTCCACTGATAGCACAGTGTCCAGAAATG	1952
Qy	641	GlnAspGluValLysThrValGlyGluGluGlnLysProGluLupProLysArgArgVal	660
Db	1953	CAAGATGAAGTCAAAACTGTGTGTGAGGAACAAAGACCCAGAGGAACCAAGCGTAGGGTG	2012
Qy	661	AspThrSerValSerTrpGluAlaLeuLysCysValGlySerSerLysLysArgAlaArg	680
Db	2013	GATACTTCAGTGTCTTGGGAGACACTGATTTGTCTCGATCATCCAGAAAGAGAGCAAGG	2072
Qy	681	LysAlaSerSerSerAspAspGluGlyGlyProArgThrLeuGlyGlyAspSerHisArg	700
Db	2073	AAGGCATCTCTTCAGATGATGAAGAGGCGCCAAAGGACACTGGGAGGGACACTCACAGA	2132
Qy	701	AlaGluGluAlaSerLysAspLysGluAlaGlyThrAspAlaValProAlaSerThrGln	720
Db	2133	CGAGAGGAGCGCCAGCAAGACAAAGAAAGCGGAAACAGACGCTGCTCTGCCACACCCAG	2192
Qy	721	GluGlnAspGlnAlaGlnGlySerSerSerProGluProAlaGlySerProSerGluGly	740
Db	2193	GAGCAGACCAAGCGCAGGAAGTTCTCTACCAGCCAGCGGAAGCCCTTCCGAAGGG	2252
Qy	741	GluGlyValSerThrTrpGluSerPheLysArgLeuValThrProArgLysLysSerLys	760
Db	2253	GAAGGTGTCTCCACTTGGGAGTCATTTAAAGATTAGTCACTCCAAAGAAAAATCCAG	2312
Qy	761	SerLysLeuGluGluLysAlaGluAspSerSerValGluGlnLeuSerThrGluLleGlu	780
Db	2313	TCAAACTCGAAGAGAAAGCCGAAGACTCTAGTGTAGAGCAGTGTGTCCACTGAGATCGAA	2372
Qy	781	ProSerArgGluGluSerTrpValSerIleLysLysPheIleProGlyArgArgLysLys	800
Db	2373	CCGAGTAGAAGAAATCTTGGGTTCATTTAAGAAATTCATCTCCCGAGCGCGGAAGAA	2432
Qy	801	ArgAlaAspGlyLysGlnGluGlnAlaThrValGluAspSerGlyProValGluLleAsn	820
Db	2433	AGGGCAGACGGGAAGCAAGAACCAAGCCACTGTGGAAGACTCAGGGCCAGTGGAGATAAT	2492
Qy	821	GluAspAspProAsnValProAlaValValProLeuSerGluTrpAsnAlaValGluArg	840
Db	2493	GAGGACGACCTTAATGTCCAGCGGTCTGTCGCTCTCTGAGTATAATGACGTGGAGAGG	2552
Qy	841	GluLysMetGluAlaGlnGlyAsnThrGluLeuProGlnLeuLeuGlyAlaValTyrVal	860
Db	2553	GAGAGATGGNAGCCCAAGGGGAATACGGAGCTGCCACGCTGCTGGGGCTGTGTAGGTG	2612
Qy	861	SerGluGluLeuSerLysThrLeuValHisThrValSerValAlaValIleAspGlyThr	880
Db	2613	TCCGAGGAGCTCAGTAAGACTCTGTGTCACACTGTGAGTGTCCAGTCTATTGATGGACC	2672
Qy	881	ArgAlaValThrSerValGluGluArgSerProSerTrpIleSerAlaSerValThrGlu	900
Db	2673	AGGGCAGTCCACAGTGTCAAGAGCGGTCTCTTCTGTGATATCCGCTTCCGTAAACAGAA	2732

Qy	901	ProLeuGluHisThrAlaGlyGluAlaMetProProValGluGluValThrGluLysAsp	920
Db	2733	CCTCTTGAACACACAGCGGAGAACCATGCGACCTGTTGAAGAGGTCACTGAAAAAGAC	2792
Qy	921	IleIleAlaGluGluThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHis	940
Db	2793	ATCATTCGAGAAGAAACTCTGTGCTCACCCAGAGCTTACCAGAGGGTAAAGATGCCCAT	2852
Qy	941	AspAspMetValThrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThr	960
Db	2853	GACGACATGTCACCCAGTGAAGTGGATTTCACCTCAGAAGCTGTGACAGCCACAGAGACC	2912
Qy	961	SerGluAlaLeuArgThrGluGluValThrGluAlaSerGlyAlaGluGluThrThrAsp	980
Db	2913	TCAGAGGCTCTCCGTACTGAAGAAGTTACCGAAGCATCGGGGGCCGAAGAGACCACAGAC	2972
Qy	981	MetValSerAlaValSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrPro	1000
Db	2973	ATGCTGTCCCGAGTTTCCCGAGCTGACTGACTCCCCAGACACACAGAGAGAACCCACCCCA	3032
Qy	1001	ValGlnGluValGluSerGlyValLeuAspThrGluGluGluGluAArgGlnThrGlnAla	1020
Db	3033	GTTCCAGAGGTAGAGAGTGGTGTGTAGATACAGAAGAAGAGAGGCCCGACCCAGAGCC	3092
Qy	1021	IleLeuGlnAlaValAlaAspLysValLysGluGluSerGlnValProAlaThrGlnThr	1040
Db	3093	ATCTCTCAAGCCCTTCGACACAAGGTGAAGAGGAGTCCCAGGTGCTGTCAACCCAGACT	3152
Qy	1041	ValGlnArgThrGlySerLysAlaLeuGluLysValGluGluValGluGluAspSerGlu	1060
Db	3153	GTGCAGAGAACCGGGTCAAAGACACTTGGAGAAAGTTTGAAGAGGTAGAGAGGACTCCGAA	3212
Qy	1061	ValLeuAlaSerGluLysGluLysAspValMetProLysGlyProValGlnGluAlaGly	1080
Db	3213	GTGCTGCTTCGGAGAAAGAGAAGACGTATTATGCCGAAGAGACCCGTGCAGAAAGCTGGAA	3272
Qy	1081	AlaGluHisLeuAlaGlnGlySerGluThrGlyGlnAlaThrProGluSerLeuGluVal	1100
Db	3273	GCTGAGCATCTTCACACAGGCTCTGAGACTGACAGGCTACTTCCAGAGAGCCTTTGAAGTT	3332
Qy	1101	ProGluValThrAlaAspValAspHisValAlaThrCysGlnValIleLysLeuGlnGln	1120
Db	3333	CCTGAAGCTCAGCGCACATGTAGACCATGTGCGCCAGCTGCCAGGTTTATCAAGCTCCAGCAG	3392
Qy	1121	LeuMetGluGlnAlaValAlaProGluSerSerGluThrLeuThrAspSerGluThrAsn	1140
Db	3393	CTGATGAACACAGCCCTGCGCCCTGAGTCAATCCGAACCTTTGACAGACACAGTGACACAAAT	3452
Qy	1141	GlySerThrProLeuAlaAspSerAspThrAlaAspGlyThrGlnGlnAspGluThrIle	1160
Db	3453	GGAGCACTCCCTTAGCAGATTCAGACACTGCAGATGGGACACACAGCAAGATCAAAACCATT	3512
Qy	1161	AspSerGlnAspSerLysAlaThrAlaAlaValArgGlnSerGlnValThrGluGluGlu	1180
Db	3513	GACAGCCAGGACAGTAAAGCCACTGCAGCTGTCCAGCGACTCACAGGTCACAGAGAAGAGAG	3572
Qy	1181	AlaAlaThrAlaGlnLysGluGluProSerThrLeuProAsnAsnValProAlaGlnGlu	1200
Db	3573	GCGGCTACTCTCAGAAAGAGAGAGCCTTCGACACTACCTAATAATGTTTCCAGCCCGCAGAA	3632
Qy	1201	GluHisGlyGluGluProGlyArgAspValLeuGluProThrGlnGlnGluLeuThrAla	1220
Db	3633	GAACATGGGAGAGAACGAGGAGAGATGTTCTTGAACTTACACAGCAAGAGCTTACTTGCT	3692
Qy	1221	AlaAlaValProValLeuAlaLysThrGluValGlyGlnGluGlyValAspTrpLeu	1240
Db	3693	GCAGCCGTGCCCTTCTGCAAAAGACTGAGGTGGGTCAAGAGGGGTGAGTTGACTGGTTG	3752
Qy	1241	AspGlyGluLysValLysGluGlnGlnGluValPheValHisSerGlyProAsnSerGln	1260
Db	3753	GATGGAGAAAAGTCAAAGAGAACGAGAGGTGTTGTATCACTCTTGAGCCCAACAGTCAA	3812


```

QY 1261 LysAlaAlaAspValThrThrAspSerGluValMetGlyValAlaGlyCysGlnGluLys 1280
Db 3813 AAGGCTGCTGAGTGTACATATGACAGTGAAGTGGAGTGGCGGGGTGACGAGAAAG 3872
QY 1281 GtuserThGluValGlnSerLeuSerLeuGluGluGluValMetGluThrAspValGlu 1300
Db 3873 GAGAGTACTGAAAGTGCAGAGTCTTACCTCGAGAGAGAGAGATGGAACCTGACCTGGA 3932
QY 1301 LysGluLysArgGluThrLysProGluGlnValSerGluGluGlyGluGlnGluThrAla 1320
Db 3933 AAGGAGAAAAAGGAGACAAAGCAGAGCAAGTGAAGAAAGTGAAGCAGGAAACAGCC 3992
QY 1321 AlaProGluHisGluGlyThrTyrgLysProValLeuThrLeuAspMetProSerSer 1340
Db 3993 GCTCTGACATGAAAGAACTTACGGAAAGCCAGTCTTACACTTGAACATGCCCCAGCTCA 4052
QY 1341 GluArgGlyLysAlaLeuGlySerLeuGlyGlySerProSerLeuProAspGlnAspLys 1360
Db 4053 GAGAGGGGAGAGGACCTGGAAAGCCTTGAAGAGAGCCCTTCTCTCCAGACCAAGACAA 4112
QY 1361 AlaGlyCysIleGluValGlnValGlnSerLeuAspThrThrValThrGlnThrAlaGlu 1380
Db 4113 GCAGGTTGATAGAGGTTCAGAGTCAAGCCTGGAACAAAGTCACTCAACAGCAGAGAA 4172
QY 1381 AlaValGluLysValIleGluThrValValIleSerGluThrGlyGluSerProGluCys 1400
Db 4173 GCTGTGAAAGGTCATGAACAGGTTGATTTCCAGAGACAGGTGAAGATCCAGAGTGT 4232
QY 1401 ValGluAlaHisLeuLeuProAlaGluLysSerSerAlaThrGlyGlyIleStrpThrLeu 1420
Db 4233 GTAGGTGCACACTTATTACAGCTGGAAGAGTCTTCTGCAAGGGGTGGCCACTGACTCTT 4292
QY 1421 GlnHisAlaGluAspThrValProLeuGlyProGluSerGlnAlaGluSerIleProIle 1440
Db 4293 CAGCATGACAGAGACACGGTACCCCTGGGGCCCTGAGTCTCAGGCAAAATCCATCCCAATC 4352
QY 1441 IleValThrProAlaProGluSerThrLeuHisProAspLeuGlnGlyGluIleSerAla 1460
Db 4353 ATAGTAACTTCCTGCTCTCTGAAACACCTTACATCTGACCTTAAAGAGAAATAAACGGA 4412
QY 1461 SerGlnArgGluAspSerGluGluGluAspLysProAspAlaGlyProAspAlaAspGly 1480
Db 4413 TCCACAGAGAGGATCAGAGAGAGAGAGAGACACAGATCTGCTGCTGATGCTGACGGC 4472
QY 1481 LysGluSerThrAlaIleGluLysValLeuLysAlaGluProGluIleLeuGluLeuGlu 1500
Db 4473 AAGGAGAGTACAGCAATCGAAAAAGTCTCAAGGCTGAACCTGAGATCTCGAACTTGAG 4532
QY 1501 SerLysSerAsnLysIleValLeuAsnValIleGlnThrAlaValAspGlnPheAlaArg 1520
Db 4533 AGTAAAGACAAACAAAGTGTGCTGAACGTCATTCACAGACCCGTTGACCGTTCGACGT 4592
QY 1521 ThrGluThrAlaProGluThrHisAlaTyraSpSerGlnThrGlnValProAlaCysArg 1540
Db 4593 ACAGAAACAGCCCCCAAACTCTGCTTATGATTCACAGCCAGGTCTCTGATGACAG 4652
QY 1541 LeuAspSerArgGluProAsnArgCysTrpThrLysMetLysAspAlaLysMetLysHis 1560
Db 4653 CTGACAGACAGGAGCCCAACAGATCTGGAACAAATGAAGATCCCAAGATGAAGAACAC 4712
QY 1561 ProValProGlnProArgGluAspLeuGlnValLeuThrValLeuGluValaTrpAlaGln 1580
Db 4713 CCGATCCCGAGCCACAGAGAGGACTTGCAAGTCTTGAACCGTTTGAGGAGATGGGCTCAG 4772
QY 1581 ProArgLysCysLeuProArgLeuGlnLeuLysAlaProValSerLys 1596
Db 4773 CCTCGAAATGCTTCCGCCCTTGCAAGTGAAGCCGCCGATGCAAG 4820

```

RESULT 5
 PCT-US02-16173A-2063
 ; Sequence 2063. Application PC/TUS0216173A
 ; GENERAL INFORMATION:
 ; APPLICANT: Mendrick, Donna

```

; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Higgs, Brandon
; APPLICANT: Castle, Arthur
; APPLICANT: Elashoff, Michael
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5089-MO
; CURRENT APPLICATION NUMBER: PCT/US02/16173A
; PRIOR FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: US 60/292,335
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/297,523
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,925
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,810
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/303,807
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/303,808
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/315,047
; PRIOR FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: US 60/324,928
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/330,867
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: US 60/330,462
; PRIOR FILING DATE: 2001-10-22
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2221
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2063
; LENGTH: 5236
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. NM_057103
PCT-US02-16173A-2063

Alignment Scores:
Pred. No.: 0 Length: 5236
Score: 8030.00 Matches: 1589
Percent Similarity: 99.75% Conservative: 3
Best Local Similarity: 99.56% Mismatches: 4
Query Match: 99.47% Indels: 0
DB: 1 Gaps: 0

US-09-902-432-4 (1-1596) x PCT-US02-16173A-2063 (1-5236)
QY 1 MetGlyAlaGlySerSerThrGluGlnArgSerProGluGlnProAlaGlySerAspThr 20
Db 33 ATGGCGGAGGAGGAGTTCACAGCAGAGCGAGAGCCCGGAGCGGAGCGACACAG 92
QY 21 ProSerGluLeuValLeuSerGlyHisGlyProAlaAlaGluAlaSerGlyValaAlaGly 40
Db 93 CCGACCGAGCTGCTCTCAATGAGGCGCATGAGGCCCGGAGCTGAAAGCTTCGGAGACACTGA 152
QY 41 AspProAlaAspAlaAspProAlaThrLysLeuProGlnLysAsnGlyGluLeuSerSer 60
Db 153 GAGCCCGCGGAGGAGCCCGGCCACCAAGCTCCACAGAAAGATGGCAGCTGTCTTCT 212
QY 61 ValaAsnGlyValaIleGluGlnGlyAspValHisValGlnGluGluAsnGlnGluGln 80
Db 213 GTCAACGGGCTGAGCAACAGAGATGTCCATGTCCAAAGAGAAACACAGAGGGGAG 272
QY 81 GluGluGluValaValaAspGluAspValaGlyGlnArgGluSerGluAspValaArgLysLys 100
Db 273 GAGGAAGAGTGTGATGAGATGTGACACGACGAGTCAAGATGTGAGAGAAAAA 332
QY 101 AspArgValaGluGluMetAlaAlaAsnSerThrAlaValaGluAspIleThrLysAspGly 120
Db 333 GACCGAGTTGAAGAAATGGCGCCCACTCCACAGCTGTGAAAGATATCACAAAGATGG 392

```

QY 121 GlnGluThrSerGluIlelleGluGlnIleProAlaSerGluAenAenValGluGlu 140
DB 393 CAGAGAGAGACATCAGAAATAATTTGAACAGATCCCTCTTCAGAAAAAATAATGTGGAAGAA 452
QY 141 MetValGlnProAlaGluSerGlnAlaAenAspValGlyPheLysValPheLysPhe 160
DB 453 ATGGTACAGCTGCTGAGTCCAGGCTTAATGATGTTGGCTTCAAGAAAGTATTAAATTT 512
QY 161 ValGlyPheLysPheThrValLysLysAspLysAsnGluLysSerAspThrValGlnLeu 180
DB 513 GTTGGTTTAAATTCACGGTGAAGAGGATAAAAAATGAAAGTCAGATACTGTCCAACTA 572
QY 181 LeuThrValLysLysAspGluGlyAlaGluAlaSerValGlyAlaGlyAspHis 200
DB 573 CTCACGTGTCAAGAAAGGATGAAGCGCAAGGGCGAAGCCCTCTGTCCGAGCTGGAGACCAC 632
QY 201 GlnGluProSerValGluThrAlaValGlyLysSerAlaSerLysLysGluSerGluLys 220
DB 633 CAGAGCCCAAGTGTGGAGACTGCCGTGGAGAGTCAGCATCCAAAGAAAGTGAAGCTGAAG 692
QY 221 GlnSerThrGluLysGlnGluGlyThrLeuLysGlnGluGlnSerSerThrGluIlePro 240
DB 693 CAATCCACAGAAAGCAAGAGGACACCTTGAAGCAAGAACAGAGCAGCACAGAAATCCCC 752
QY 241 LeuGlnAlaGluSerAspGlnAlaAlaGluGluAlaLysAspGluGlyGluGluLys 260
DB 753 CTTCAAGCCCAATCTGATCAAGCGCTGAGCAAGAACCCAAAGATGAAGAGAGAAAAA 812
QY 261 GlnGluLysGluProThrLysSerProGluSerProGluSerProValAenSerGluThr 280
DB 813 CAAGAAAGAGGCCCAACCAAGTCCCCAAGATCCCGAGCAGCCAGCCAGTCAACAGTGAACA 872
QY 281 ThrSerSerPheLysLysPheThrHisGlyTyrAlaGlyTyrArgLysLysThrSer 300
DB 873 ACATCTCTCTTCAAGAAAGTTCTCACTACAGGTTGGCCCGCTGGCCCAAGAGACCAGC 932
QY 301 PheLysLysSerLysGluAspAspLeuGluThrAlaGluLysArgLysGluGlnGluAla 320
DB 933 TTCAAGAAATCAAAAGAGGATGATCTGAAATCTGCCAGAAAGAGAAAGAGAGGCA 992
QY 321 GluLysValAspGluGluGluLysGluLysThrGluProAlaSerGluGlnGluPro 340
DB 993 GAAAAAGTAGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1052
QY 341 AlaGluAspThrAspGlnAlaArgLeuSerAlaAspTyrGluLysValGluLeuProLeu 360
DB 1053 GCAGAGACACAGACCAGGCGAGGTTGTACAGACTACAGAGAGGTTGGAGCTGCCCTTTG 1112
QY 361 GluAspGlnValGlyAspLeuGluAlaSerSerGluGluLysCysAlaProLeuAlaThr 380
DB 1113 GAAGACCAGGTTGGTACCTGGAGGCATCTGTACAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1172
QY 381 GluValPheAspGluLysMetGluAlaHisGlnGluValValAlaGluValHisValSer 400
DB 1173 GAAGTGTGTTGATGAGAGATGGAAGGCCACCAAGAACTGTTGCAGAGGTTCCACGTGAGC 1232
QY 401 ThrValGluLysThrGluGluGlnGlnGlyGlyGlyGlyGlyGlyGlyGlyGlyValVal 420
DB 1233 ACCGTGAGAGAGACAGAGGAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1292
QY 421 ValGluGlyThrGlyGluSerLeuProProGluLysLeuAlaGluProGlnGluValPro 440
DB 1293 GTAGAGAGACAGAGAGATCTTGGCCCTGAGAACTGGCTGAGCCCAAGAGGTTCCCC 1352
QY 441 GlnGluAlaGluProAlaGluGluLeuMetLysSerArgGluMetCysValSerGlyGly 460
DB 1353 CAGGAAGCTGAGCTGTGAGGAGCTGATGAAGAGCAGAGAGATGTGTCTCTGGAGGA 1412
QY 461 AspHisThrGlnLeuThrAspLeuSerProGluGluLysThrLeuProLysHisProGlu 480
DB 1413 GACCACACTCACTGACAGACCTAAGTCTCTGAAGAGAGAGGCTGCCCAACACCCAGAA 1472

QY 481 GlyIleValSerGluValGluMetLeuSerSerGlnGluArgIleLysValGlnGlySer 500
DB 1473 GGCATTGTTCAGTTCAGTGGAGATGCTGCTCTCAGAAAAGAAATCAAGGTACAGGGAAGT 1532
QY 501 ProLeuLysLysLeuPheSerSerSerGlyLeuLysLysLeuSerGlyLysLysGlnLys 520
DB 1533 CCCTTGAGAGAACTCTTCAGTAGCTCAGGCTTAAAGAAAGCTGTCTGGGAAGACGAGAAG 1592
QY 521 GlyLysArgGlyGlyGlyGlyAspGluGluProGlyGluTyrGlnHisIleThrGlu 540
DB 1593 GGGAAACGAGGAGGTGGCGAGACGAAAGAGCTGGAGAAATACCAACACATTCACACCGAA 1652
QY 541 SerProGluSerAlaAspGluGlnLysGlyGluSerSerAlaSerSerProGluGluPro 560
DB 1653 TCCCCAGAGAGTCTGATGAGCAGAAAGGAGAGAGCTCTCGCTCGTCCCCCGAGGAGCCT 1712
QY 561 GluGluThrThrCysLeuGluLysGlyProLeuGluAlaProGlnAspGlyGluAlaGlu 580
DB 1713 GAGGAGACCACGTGTCTGGAGAAAGGCCCTGGAAAGACCCAGGATGGGAAAGCTGAG 1772
QY 581 GluGlyThrThrSerAspGlyGluLysLysArgGluGlyIleThrProTrpAlaSerPhe 600
DB 1773 GAAGGAACCTACTTCCGATGGAGAGAGAGAGAGAGGATCACTCCCTGGGCATCTCTTC 1832
QY 601 LysLysMetValThrProLysLysArgValArgArgProSerSerGluSerAspLysGluGlu 620
DB 1833 AAAAAGATGTTGACACCCCAAGAAACGGGTCCGAAGACCTTCTGAGAGTGACAAAGGAGGA 1892
QY 621 GluLeuGluLysValLysSerAlaThrLeuSerSerThrAspSerThrValSerGluMet 640
DB 1893 GAGCTGGAGAAAGGTCAAGAGCGCCACTTGTCTCTCACTGATAGCACAGTGTCCAGAAATG 1952
QY 641 GlnAspGluValLysThrValGlyGluGluGlnLysProGluProLysArgArgVal 660
DB 1953 CAAGATGAAGTCAAACTGTTGGTGAGGAACAAAGCCAGAGAAACCAAGACCTAGGGTG 2012
QY 661 AspThrSerValSerTrpGluAlaLeuLysCysValGlySerSerLysLysArgAlaArg 680
DB 2013 GATACCTCAGTGTCTTGGGAAGCACTGATTTGTGTCGATCATCCAAAGAGAGAGCAAGG 2072
QY 681 LysAlaSerSerSerAspAspGluGlyGlyProArgThrLeuGlyGlyAspSerHisArg 700
DB 2073 AAGGCATCTCTTCAGATGATGAAGGAGGCGCCAAAGACACTGGGAGGGGACAGTCCAGA 2132
QY 701 AlaGluGluAlaSerLysAspLysGluAlaGlyThrAspAlaValProAlaSerThrGln 720
DB 2133 GCAGAGAGGCCAGCAACAGCAAGAGCCGNAACAGACGCTGTTCTGCCAGCACCCAG 2192
QY 721 GluGlnAspGlnAlaGlnGlySerSerSerProGluProAlaGlySerProSerGluGly 740
DB 2193 GAGCAGGACCAAGCGCAAGGAAGTTCTCCACCCGAGCCAGCGGGAAGCCCTTCCGAAGGG 2252
QY 741 GluGlyValSerThrTrpGluSerPheLysArgLeuValThrProArgLysLysSerLys 760
DB 2253 GAAGTGTCTCCACTTGGGAGTCATTAAAGAGATTAGTCACCTCCAAGAAAAAATCCAAG 2312
QY 761 SerLysLeuGluGluLysAlaGluAspSerSerValGluGlnLeuSerThrGluIleGlu 780
DB 2313 TCAAAACCTGGAAGAGAAACCCGAAGACTCTAGTGTAGACAGATTGTCTACTGAGATCGAA 2372
QY 781 ProSerArgGluGluSerTrpValSerLysLysPheIleProGlyArgArgLysLys 800
DB 2373 CCGAGTAGAGAGAAATCTTGGGTTTCCATTAAAGAAATTCATCCCGGACGGCGGAGAGAA 2432
QY 801 ArgAlaAspGlyLysGlnGluGlnAlaThrValGluAspSerGlyProValGluIleAsn 820
DB 2433 AGGGCAGATGGGAAGCAAGAACCAAGCCACTGTGGAAGACTCAGGGCCAGTGGAGATAAT 2492
QY 821 GluAspAspProAsnValProAlaValValProLeuSerGluTyrAenAlaValGluArg 840
DB 2493 GAGGAGACCCCTGATGTCCAGCCGTCGCTCTGCTGAGTATGATGTCAGTGGAGAGG 2552
QY 841 GluLysMetGluAlaGlnGlyAsnThrGluLeuProGlnLeuGlyAlaValTyrVal 860

Db 2553 GAGAAAGATGGAAGCCAGGAGGAATGGGAGCTCCCGACCTGCTGGGGCTGTGATCGTG 2612
 QY 861 SGGJUGLULeuSerIysThrLeuValHisThrValSerValAlaVal11LeuApGIYThr 880
 Db 2613 TCCGAGAGACTCAGTAAAGACTGTGTCCACACTGTAGTGTCCAGTCACTTGTGGAGACC 2672
 QY 881 ArgAlaVal1ThrSerValIGLUglUatGserProSerTTrp1IleSer1AsSerVal1ThrGIU 900
 Db 2673 AAGGAGCTACCAAGTGTCCAAAGAGCGGTCTCTTGTGATATCCGCTCCGTAACAGAA 2732
 QY 901 ProLeuGIUHI:sthrAlaGIUglUaMetProProValIGLUglUaVal1ThrGIUlyAsp 920
 Db 2733 CCTCTTGAACACACAGCGGAGAGAGCCATGCCCTGTGAAAGAGTCACTGAAAAAGAC 2792
 QY 921 IleIleAlaGIUglUuThrProVal1LeuThrGIUThrLeuProGIUglUyAspAlaHis 940
 Db 2793 ATCATTTGCAAGAAACTCTGTGTCTCACCCAGACGTTACAGAGGGGTAAAGATGCCAT 2852
 QY 941 AspAspMetVal1ThrSerGIUVal1AspPheThrSerGIUAlaVal1ThrAlaThrGIUThr 960
 Db 2853 GACGACATGTGACACCGATGAGATGATTTCACTCAGAAAGCTGTGACAGCCACAGAGACC 2912
 QY 961 SerGIUAlaLeuArgThrGIUglUuVal1ThrGIUAlaSerGIUVal1AGLUglUuThrThrAsp 980
 Db 2913 TCGAGGCTCTCGTACTGAGAGATTACCGAAGCATCGGGGCGCGAAGAGACCAAGAC 2972
 QY 981 MetValSerAlaValSerGIULeuThrAspSerProAspThrThrGIUglUa1aThrPro 1000
 Db 2973 ATGGTGTCCGCAATTTCCACAGTGACTGACTCCCGACAGACCAAGAGGAGCCACCCCA 3032
 QY 1001 ValGIUglUuValIGLUserGIUVal1LeuAspThrGIUglUuGIUuArgGIUThrGIUAla 1020
 Db 3033 GTTCAGAGAGTGAAGAGTGTGTGATACAGAAAGAGAGAGGCCAGAGAGCGGCC 3092
 QY 1021 IleLeuGIUAlaValAlaAspLysVal1LysGIUglUuSerGIUVal1ProAlaThrGIUThr 1040
 Db 3093 ATCTTCACAGCCCTTCCAGACAGGTGAAGAGAGAGTCCAGGTGCTGCAACCCAGACT 3152
 QY 1041 ValGIUArgThrGIUySerLysAlaLeuGIUlyVal1GIUglUuValIGLUuAspSerGIU 1060
 Db 3153 GTGCAGAGAGACGGGGTCAAAAGCATCGAAGAGTTGAGAGAGTTAAGAGAGACTCCGAA 3212
 QY 1061 ValLeuAlaSerGIUlySGLUyAspVal1MetProLysGIUProValGIUglUa1aGIU 1080
 Db 3213 GTGCTGGCTTCGAGAAAGAGAGAGCGTTATGCCAAAGAGACCCGTGACGAGAGCTGGA 3272
 QY 1081 AlaGIUHisLeuAlaGIUngIySerGIUThrGIUAla1aThrProGIULeuGIUVal1 1100
 Db 3273 GCTGAGCATCTTGACAGAGGCTCTGAGACTGACAGAGCTACTCCAGAGACCTTGAAGTT 3332
 QY 1101 ProGIUVal1ThrAlaAspVal1AspHisVal1a1aThrCysGIUngIVal1LeuLysLeuGIUngI 1120
 Db 3333 CCTGAAGTCAAGCGCAATGTAGACATGTGTGCCAGGTGCAAGTTTCAAGCTCCAGCGAG 3392
 QY 1121 LeuMetGIUglUAlaVal1a1aProGIUySerGIUThrGIUThrAspSerGIUThrAsn 1140
 Db 3393 CTGATGGAACAGGCGCTGGCGCTCTGAGTATCCGAAACCTTGACACACAGTGAAGACAAAT 3452
 QY 1141 GlySerThrProLeuAlaAspSerAspThrAlaAspGIUThrGIUngIuAspGIUThrIle 1160
 Db 3453 GGAAGAGATCTCTTACACAGATTCAGACACTGACAGTGGGACACAGCAAGATGAAACCAT 3512
 QY 1161 AspSerGIUAspSerLysAlaThrAlaAlaVala1aGIUngIuSerGIUVal1ThrGIUglUglU 1180
 Db 3513 GACAGCCAGAGACAGTAAAGCACTGCAAGTGTCAAGGACATCAAGGTCAACAGAAAGAG 3572
 QY 1181 AlaAlaThrAlaGIUlySGLUglUProSerThrLeuProAsnAsnVal1ProAlaGIUngIu 1200
 Db 3573 GCGGCTACTGCTCAGAAAGAGAGCCCTTGACACTACTAATATGTTCAGGCCCGAGAA 3632
 QY 1201 GIUHisGIUglUglUProGIUArgAspVal1LeuGIUProThrGIUngIuLeuThrAla 1220

Db 3633 GAACATGGGGAAGAACCAAGAGATGTTCTTGAACCTTACACAGCAAGACTTACTGCT 3692
 QY 1221 AlaAlaVal1ProVal1LeuAla1a1ySerThrGIUVal1GIUngIuGIUglUVal1AspTrpLeu 1240
 Db 3693 GCAGCCGTGCCCTTCTTGCGAAAGACTGAGTGGGTCAAGAGGGGTGAGTGTGACTGTGG 3752
 QY 1241 AspGIUglUyVal1ySGLUglUngIuVal1PheVal1HisSerGIUProAsnSerGIU 1260
 Db 3753 GATGAGAAAAAGTCAAAAGAAAGACAGAGGTGTTTGTACACTCTGGACCCCAAGCTAA 3812
 QY 1261 LysAlaAlaAspVal1ThrTyAspSerGIUVal1MetGIUVala1aGIUySGLUngIuLys 1280
 Db 3813 AAGCTCTGATGTGACATATGACAGTGAAGTGAAGTGGAGTGGCGGGGTGCAGGAAAG 3872
 QY 1281 GIUThrThrGIUVal1GIUngIuSerLeuSerLeuGIUglUglUyGIUuMetGIUThrAspVal1GIU 1300
 Db 3873 GAGACTACTGAAGTGCACAGTCTTACCTGGAGGAGGAGATGGAATGCAACTGACTTAA 3932
 QY 1301 LysGIUlyAspArgGIUThrLysProGIUngIuVal1SerGIUngIuGIUglUngIuThrAla 1320
 Db 3933 AAGGAGAAAAAGGAGACAAAGCCAGAGCAAGTGAAGTGAAGAGTGAACAGAGAAACAGCC 3992
 QY 1321 AlaProGIUHI:SGLUglUyThrTyGIUyAspProVal1LeuThrLeuAspMetProSerSer 1340
 Db 3993 GCTCTGAGACATGAAGAGAACCTACGGGAAGCCAGTCTGACACTTGAACATGCCAGCTCA 4052
 QY 1341 GIUArgGIUyVal1LeuGIUySerLeuGIUyGIUySerProSerLeuProAspGIUAspLys 1360
 Db 4053 GAGAGGGGGAAGGCACCTGGGAAGCCTTGAGAGAGCCCTTCTCTCCAGACCAACAGAAA 4112
 QY 1361 AlaGIUyCys1IleGIUngIuVal1GIUngIuSerLeuAspTrpThrVal1ThrGIUThrAlaGIU 1380
 Db 4113 GCAGGTGTATAGAGTTCAGATTCAAGCTTGACACACAGTACTCAACACACAGAA 4172
 QY 1381 AlaVal1GIUlyVal1IleGIUThrVal1IleSerGIUThrGIUyGIUySerProGIUyCys 1400
 Db 4173 GCTGTGAAAAAGTCAATGAAGACGGTTGTGATTTCAAGACAGAGTGAAGATCCAGAGGT 4232
 QY 1401 ValGIUyAlaHisLeuLeuProAlaGIUlySerSerAla1aThrGIUyLys1IleTrpThrLeu 1420
 Db 4233 GTAGGTGCACACTTATTTCACAGCTGAGAAAGTCTCTGCAACGGGGCCACTGCACTCTT 4292
 QY 1421 GIUHisAlaGIUAspThrVal1ProLeuGIUProGIUySerGIUAlaGIUser1IleProIle 1440
 Db 4293 CAGCATGCAAGAGACAGGTACCTCGGGGCTGTGATCTCAGGCAAGATCCATCCCAATC 4352
 QY 1441 IleVal1ThrProAlaProGIUySerThrLeuHisProAspLeuGIUngIyGIU1IleSerAla 1460
 Db 4353 ATAGTAACTCCGTCTCTGAAGAACCCCTACATCTCTGACCTTACAGAGAGAAATTAAGCGCA 4412
 QY 1461 SerGIUArgGIUArgSerGIUglUglUAspLysProAspAlaGIUProAspAlaAspGIU 1480
 Db 4413 TCCCAAGAGAGCGATACAGAGGAGAGACAAAGCAGATGTGTGTCTGATGCTGACAGCGC 4472
 QY 1481 LysGIUThrAla1IleGIUlySerVal1LeuLysAlaGIUProGIU1IleLeuGIUngIu 1500
 Db 4473 AAGGAGAGTACAGCAATGAAAGATCTCTCAAGGCTGAAACCTGAAGATCTTGAACTTGA 4532
 QY 1501 SerLysSerAsnLys1IleVal1LeuAsnVal1IleGIUThrAlaAlaAspGIUngIuAlaArg 1520
 Db 4533 AGTAAGAGCAAAAGATTTGTGTGAACGTCAATTCAGACAGCCGTGTGACCAAGTTCACAGT 4592
 QY 1521 ThrGIUThrAlaProGIUThrHisAlaTyAspSerGIUThrGIUngIuVal1ProAlaCysArg 1540
 Db 4593 ACAGAAACAGCCCCGAAACCTCATGCTTATGATTCACAGACCCCAAGTGTCTGCAATGCGAG 4652
 QY 1541 LeuAspSerArgGIUProAsnArgCysTrpThrLysMetLysAspAlaLysMetLysHis 1560
 Db 4653 GCTGACAGAGGAGGCCAAACAGATGTGAGACAAAATGAAGATTGTCACAGATGAAGAAC 4712
 QY 1561 ProVal1ProGIUProArgGIUAspLeuGIUngIuVal1LeuThrVal1LeuGIUAla1aThrAlaGIU 1580
 Db 4713 CCAAGTCCCGCACCCAGAGAGGACTTGCAGAGTCTTGACCGTTCTGAGAGCAAGGGCTCAG 4772

QY 1581 ProArgLysCysLeuProArgLeuGlnLeuAlaProValSerLys 1596
 Db 4773 CTCGGAATGCTTGGCGCGTTCAGTGAAGAGCGCGGTGCAAG 4820

RESULT 6

US-10-191-803-279
 ; Sequence 279, Application US/10191803
 ; GENERAL INFORMATION:
 ; APPLICANT: MENDRICK, Donna
 ; APPLICANT: PORTER, Mark
 ; APPLICANT: JOHNSON, Kory
 ; APPLICANT: HIGGS, Brandon
 ; APPLICANT: CASTLE, Arthur
 ; APPLICANT: ELASHOFF, Michael
 ; TITLE OF INVENTION: CardioToxin Molecular Toxicology Modeling
 ; FILE REFERENCE: 44921-5090US
 ; CURRENT APPLICATION NUMBER: US/10/191,803
 ; CURRENT FILING DATE: 2002-07-10
 ; PRIOR APPLICATION NUMBER: US 60/303,819
 ; PRIOR FILING DATE: 2001-07-10
 ; PRIOR APPLICATION NUMBER: US 60/305,623
 ; PRIOR FILING DATE: 2001-07-17
 ; PRIOR APPLICATION NUMBER: US 60/369,351
 ; PRIOR FILING DATE: 2002-04-03
 ; PRIOR APPLICATION NUMBER: US 60/377,611
 ; PRIOR FILING DATE: 2002-05-06
 ; NUMBER OF SEQ ID NOS: 1140
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 279
 ; LENGTH: 5236
 ; TYPE: DNA
 ; ORGANISM: Rattus norvegicus
 ; FEATURE:
 ; OTHER INFORMATION: Genbank Accession No. NM_057103
 US-10-191-803-279

Alignment Scores:

Pred. No.: 0 Length: 5236
 Score: 8030.00 Matches: 1589
 Percent Similarity: 99.75% Conservative: 3
 Best Local Similarity: 99.56% Mismatches: 4
 Query Match: 99.47% Indels: 0
 DB: 41 Gaps: 0
 US-09-902-432-4 (1-1596) x US-10-191-803-279 (1-5236)

QY 1 MetGlyAlaGlySerSerThrGluGlnArgSerProGluGlnProAlaGlySerAspThr 20
 Db 33 ATGGGCGCAGCGAGTTCACCGAGCAGCGAGCGCGCGCGCGCGCGCGCGCGCGCG 92
 QY 21 ProSerGluLeuValLeuSerGlyHisGlyProAlaAlaGluAlaSerGlyAlaAlaGly 40
 Db 93 CCAGCGAGCTGGTGTCTAGTGGCCATGGGCGCGCGCGCGCGCGCGCGCGCGCGCG 152
 QY 41 AspProAlaAspAlaAspProAlaThrLysLeuProGlnLysAsnGlyGlnLeuSerSer 60
 Db 153 GACCCCGCGAGCGGACCCCGCCACCAAGCTCCACAGAAAGTGGCCAGCTGTCTTCT 212
 QY 61 ValAsnGlyValAlaGluGlnGlyAspValHisValGlnGluGlnGluGlyGln 80
 Db 213 GTCAACCGCGTAGCTGAACAGGAGATGTCCATGTCCAAGAGGAAACCAGAGGGGCG 272
 QY 81 GluGluGluValValAspGluAspValGlyGlnArgGluSerGluAspValArgGluLys 100
 Db 273 GAGGAAGTCTGTGTGATGAGGATGTTGGACAGCGAGTGAAGATGTGAGAGAAAA 332
 QY 101 AspArgValGluGluMetAlaAlaAsnSerThrAlaValGluAspIleThrLysAspGly 120
 Db 333 GACCGAGTTGAAGAAATGGCGGCACTCCACAGCTGTTGAGATATACAAAGGATGG 392
 QY 121 GlnGluThrSerGluIleIleGluGlnIleProAlaSerGluAsnValGluGlu 140

Db 393 CAGGAGGAGACATCAGAAAATAATTGAACAGATCCCTGCTTCAGAAAACAATGTGGAGAA 452
 QY 141 MetValGlnProAlaGluSerGlnAlaAsnAspValGlyPheLysLysValPheLysPhe 160
 Db 453 ATGGTACAGCGCTGCTGAGTCCCAGGCTAATGATGTTGGCTTCAAGAAAGTATTATAATTT 512
 QY 161 ValGlyPheLysPheThrValLysLysAspLysAsnGluLysSerAspThrValGlnLeu 180
 Db 513 GTTGGTTTTAAATTTCACGGTGAAGAGGATAAAAAATGAAAGTCAAGTACTGTCCAACTA 572
 QY 181 LeuThrValLysLysAspGluGlyGluAlaGluAlaSerValGlyAlaGlyAspHis 200
 Db 573 CTCACTGTCAAGAGGATGAAGCGGAGGGCGAAGCCCTCTGTCGAGCTCGAGACCAC 632
 QY 201 GlnGluProSerValGluThrAlaValGlyGluSerAlaSerLysGluSerGluLeuLys 220
 Db 633 CAGGAGCCCGAGTGTGAGAGTCCGTCGAGAGTCAAGCATCCAAAGAAAGTGAAGTGAAG 692
 QY 221 GlnSerThrGluLysGlnGluGlyThrLeuLysGlnGluGlnSerSerThrGluIlePro 240
 Db 693 CAATCCACAGAGAGCAAGAGCCACCTCAAGCAAGAAACAGAGCAGCAGCAAGAAATCCCC 752
 QY 241 LeuGlnAlaGluSerAspGlnAlaAlaGluGluAlaLysAspGluGlyGluGluLys 260
 Db 753 CTTCAAGCCGATCTGATCAAGCGCTGAGGAAGAACCAAGATCAAGAGGAAGAAAA 812
 QY 261 GlnGluLysGluProThrLysSerProGluSerProSerSerProValAsnSerGluThr 280
 Db 813 CAAGAGAAAGAGCCCAACCAAGTCCCCAGAAATCCCCGAGCAGCCCACTCAACACTGAGACA 872
 QY 281 ThrSerSerPheLysLysPheThrHisGlyTrpAlaGlyTrpArgLysLysThrSer 300
 Db 873 ACATCTTCTTCAAGAAAGTTCTTCACTACGGTGGCGCGCTGGCGCAAGAGACGACG 932
 QY 301 PheLysSerLysGluAspLeuGluThrAlaGluLysArgLysGluGlnGluAla 320
 Db 933 TTCAGAAATCAAAAGAGGATGATCTGGAAACTGCGCAAGAGAAAGGAGCAAGAGGCA 992
 QY 321 GluLysValAspGluGluLysGluLysThrGluProAlaSerGluGluGlnGluPro 340
 Db 993 GAAAAAGTAGACGAGAGAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1052
 QY 341 AlaGluAspThrAspGlnAlaArgLeuSerAlaAspTyrGluLysValGluLeuProLeu 360
 Db 1053 GCAGAAAGACACAGACAGCCAGGCTTGTCTAGCAGACTACAGAGAGGAGGAGGCTGAA 1112
 QY 361 GluAsnGlnValGlyAspLeuGluAlaSerSerGluGluLysCysAlaProLeuAlaThr 380
 Db 1113 GAAGACAGGTTGGTGGTGGAGGAGTGTCTGAGAGAGAGTGTCTCTTTGGCAACG 1172
 QY 381 GluValPheAspGluLysMetGluAlaHisGlnGluValValAlaGluValHisValSer 400
 Db 1173 GAAGTGTGTGATGAGAGAGTGGNAGCCCAAGAGAGTGTTCAGAGGTTCCACGTCGAGC 1232
 QY 401 ThrValGluLysThrGluGluGlnGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyVal 420
 Db 1233 ACCGTGGAGAAAGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1292
 QY 421 ValGluGlyThrGlyLysSerLeuProGluLysLeuAlaGluProGlnGluValPro 440
 Db 1293 GTAGAGGAGAACAGGAGAAATCTTTGCCCTGTGAGAACTGGCTGAGCCCGCAGGAGGTC 1352
 QY 441 GlnGluAlaGluProAlaGluGluLeuMetLysSerArgGluMetCysValSerGlyGly 460
 Db 1353 CAGGAAGCTGAGCTCTCTGAGGAGCTGTGAAGAGCAGAGAGATGTGTCTCTGGAGGA 1412
 QY 461 AspHisThrGlnLeuThrAspLeuSerProGluGluLysThrLeuProLysHisProGlu 480
 Db 1413 GACCACACTCACTGACAGACCTAAGTCTCTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 1472
 QY 481 GlyIleValSerGluValGluMetLeuSerSerGlnGluArgIleLysValGlnGlySer 500
 Db 1473 GGCATTGTGAGTGGAGATGCTGTCTCTCTCAGGAAAGAAATCAAGGTACAGGGAAGT 1532

QY	501	ProLeuylsLysLeuPheSerSerSergLysLeuLysLeuSerGlyLysLysGluLys	520
Db	1533	CCCTTGAAAGAACTCTTCAGTAGCTCAGCGCTTAAAGAACTGCTGGGAAAGAACAGAG	1592
QY	521	GlyLysAsnGlyGlyGlyLysAspGluGluProGlyGluThrGluHistiLeithrGlu	540
Db	1593	GGGAAACAGAGAGGTGGGGAGAGCAAGACCTCGAGAAATACCAACACTTCACCGAA	1652
QY	541	SerProGluSerAlaAspGluGluLysGlyLysSerAlaSerSerProGluGluPro	560
Db	1653	TCGCCAGAGAGTGTCTGATGACAAAGAGAGAGACTCTGCTGCTCCCCCGAGAGACT	1712
QY	561	GluGluThrThrCysLeuGluLysGlyProLeuGluAlaProGluAspGlyGluAlaGlu	580
Db	1713	GAGGAGACCAAGTGTCTGGAGAAAGGGCCGCTGGAGACCCAGAGATGGGAACTAG	1772
QY	581	GluGlyThrThrSerAspGlyGlyLysLysArgGlyLysIleThrProThrAlaSerPhe	600
Db	1773	GAGGAACCTACTTCCGATGGAGAAAGAAAGAGGAGATCACTCCCGGACACTTC	1832
QY	601	LysLysMetValThrProLysLysArgValAlaArgProSerGluSerAspLysGluGlu	620
Db	1833	AAAGAGATGTGACACCCAAAGAAACGGGCTCCAGACCTTCGAGAGTACAAAGAGAA	1892
QY	621	GluLeuGluLysValLysSerAlaThrLeuSerSerThrAspSerThrValSerGluMet	640
Db	1893	GAGCTGGAGAAAGTCAAGAGCGCCACTTGTCTCCACTGATGACACTGTCAAGAAATG	1952
QY	641	GluAspGluValLysThrValGlyGluGluGluLysPheGluGluProLysArgArgVal	660
Db	1953	CAAGATGAAGTCAAAACCTGTGGTGAGGAAACAAACCCAGAGAACCAAGGCTGGGTG	2012
QY	661	AspThrSerValSerThrProGluAlaLeuIleCysValGlySerSerLysLysArgAlaArg	680
Db	2013	GATACTTAGAGTCTTGGAAGACCTGATTTGTGTGATCATTCACAAAGAGACCAAGG	2072
QY	681	LysAlaSerSerSerAspAspGluGlyGlyProArgThrLeuGlyGlyAspSerHisArg	700
Db	2073	AAGGATCTCTTCAGATGATGAAGAGGGCCAAAGGACCTGGAGGGGACGTACAGAG	2132
QY	701	AlaGluGluAlaSerLysAspLysGluAlaGlyThrAspAlaValProAlaSerThrGlu	720
Db	2133	GCAGAGGAGGCGCACAAAGCAAAAGAACCGGAAACAGAGCTGTCTCGCAGCACCCAG	2192
QY	721	GluGluAspGluAlaGluGlySerSerSerProGluProAlaGlySerProSerGluGly	740
Db	2193	GAGCAGGACCAAGGCGCAAGAGAGTTCTCTCACCGAGCAGCGGAGAGCCCTTCGAAAGG	2252
QY	741	GluGlyValSerThrThrProGluSerPheLysArgLeuValThrProArgLysLysSerLys	760
Db	2253	GAAAGTGTCTCCACTTGGGAGCTCTTTAAAGATTAGTCACTCCAAAGAAAAAAATCCAG	2312
QY	761	SerLysLeuGluGluLysAlaGluAspSerSerValGluGluLeuSerThrGluIleGlu	780
Db	2313	TCAAAACTGAAAGAGAAAGCCGAAAGACTTAAAGTGTAGACACAGTTCACACTAGATCGAA	2372
QY	781	ProSerArgGluGluSerThrProValSerIleLysLysPheIleProGlyLysArgLysLys	800
Db	2373	CCGAGTAAAGAAAGATCTTGGGTTTCCATTAAAGAAATTCATCCCGGAGCGGAGAAAGAA	2432
QY	801	ArgAlaAspGlyLysGluGluGluAlaThrValGluAspSerGlyProValGluIleAsn	820
Db	2433	AGGGCAGATGGAGAACCAAGACCAAGCCTGTGAAAGACTCAAGGGCCAGTGGAGATTAAT	2492
QY	821	GluAspAspProAsnValProAlaValValProLeuSerGluThrAsnAlaValGluArg	840
Db	2493	GAGGACGACCTCGATGTGCCAGCGGTGTGCTCTGTGTGAAGTATGATCACTGAGAGAGG	2552
QY	841	GluLysMetGluAlaGluGlyAsnThrGluLeuProGluLeuLeuGlyAlaValAlaLysVal	860
Db	2553	GAGAAAGATGGAAGCCAGGGGAAATGCGGAGCTCCCACTGCTGGGGCTGTGAACGTG	2612

QY	861	SErGIuLueSerLystrhLeuValhIsthrrValSerValAlaValIleAspGlyThr	880
Db	2613	TCcAGAGAGCTcAGTAAgACTcGTGCTCCACtGTGAGTGTCCAGTATTGAAGGAGACC	2672
QY	881	ArgAlaValThrSerValGluGluArgSerProSerTrpIleSerAlaSerValThrGlu	900
Db	2673	AGGGcAGTCACcAGTGTCCAAAGcCGGTCTCTTGTGTGAATTCGGTTCCTTAACGAA	2732
QY	901	ProLeuGluIsthrrAlaGluGluAlaMetProProValGluGluValThrGluLysAsp	920
Db	2733	CCTCTTGAAcACAGAGGGGAGAAgCCATGCCACCTGTTGAAGAGGTCACTGAAGAAAGAC	2792
QY	921	IleIleLeaGluGluThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHis	940
Db	2793	ATcATTTCAGAAAGAAcTCTGTGTCCACcAGACGTTAACAGAGGTGAAGATGCCAT	2852
QY	941	AspAspMetValThrSerGluValAspPheThrSerGluAlaValAlaThrAlaThrGluThr	960
Db	2853	GAACACATGTCTACcAGTGAAGTGGATTTCACCTCGAAGCTGTACAGCCACAGAGACC	2912
QY	961	SerGluAlaLeuArgThrGluGluValThrGluAlaSerGlyAlaGluGluThrThrAsp	980
Db	2913	TCAGAGGTCTCCGTACTGAAGAGTTACCGAAGCTGGGGGGCCGAAGAGACACAGAC	2972
QY	981	MetValSerAlaValSerGluLeuThrAspSerProAspThrThrGluAlaAlaThrPro	1000
Db	2973	ATGGTGTCCGAGTTTCCcAGCTGACTGACTCCcAGACACACAGAGAAAGCCACCCCA	3032
QY	1001	ValGluGluValGluSerGlyValLeuAspThrGluGluGluArgGluThrGluAla	1020
Db	3033	GTTcAGAGGTAGAGGTGTGTGTACTAGTAAcGAAAGAAAGAGACCCGCAACCGAGCC	3092
QY	1021	IleLeuGluAlaValAlaAspLysValLysGluGluSerGluValProAlaThrGluThr	1040
Db	3093	ATCTCTCAAGCCGTGTGAGACCAAGTGAAGAGAGTCCcAGATGCTCGACACCAGACT	3152
QY	1041	ValGluArgThrGlySerLysAlaLeuGluLysValGluGluValGluLysAspSerGlu	1060
Db	3153	GTGCAGAGAAcCGGGGTCAAAAGCACTGGAGAAgGTTGAGAGGTGAAGAGACTCCGAA	3212
QY	1061	ValLeuAlaSerGluLysGluLysAspValMetProLysGlyProValGluGluAlaGly	1080
Db	3213	GTGCTGGTGTGGAGAAAGAGAGcGTTATGCCAAAGACCCGTCCAGAGATGTGGA	3272
QY	1081	AlaGluIsthrrLeuAlaGluGlySerGluThrGlyGluAlaThrProGluSerLeuGluVal	1100
Db	3273	GCTAGACATCTTGACACAGcGCTTGAGACTGGAcAGGTCTCCAGAGAGcCTTGAAGTT	3332
QY	1101	ProGluValThrAlaAspValAspHisValAlaThrCysGluValIleLysLeuGluGln	1120
Db	3333	CCTGAAGTCACGGcAGATGTAGACCATGTGTCCGACGTCGAGATTATCAAGCTCCAGAG	3392
QY	1121	LeuMetGluGluAlaValAlaProGluSerSerGluThrLeuThrAspSerGluThrAsn	1140
Db	3393	CTGTATGAAACAGcCGGTGGCCCTGTAGTATCCGAAcCTTTCACAGACAGTGAACAAT	3452
QY	1141	GlySerThrProLeuAlaAspSerAspThrAlaAspGlyThrGluAlaAspGluThrIle	1160
Db	3453	GAAAGCACTCCCTTAGcAGATTTAGACACTCGAGATGGGACACAGCAAGATGAACAATT	3512
QY	1161	AspSerGluAspSerLysAlaThrAlaAlaValArgGluSerGluValThrGluGluGlu	1180
Db	3513	GACAGCCAGACAGTAAAGcCATGcAGCTGTTCAGGcAGTCAAGGTCAcAGAAAGAGAG	3572
QY	1181	AlaAlaThrAlaGluLysGluGluProSerThrLeuProAsnAsnValProAlaGluGlu	1200
Db	3573	GGCGTCTACTGTcAGAAAGGAGcCTTGACACTACCTAAATAAATGTTCCAGCCcAGAA	3632
QY	1201	GluIsthrrGluGluProGlyArgAspValLeuGluProThrGluGluGluThrAla	1220
Db	3633	GAACATGGGAAAGAAcCGAGGAAGATGTTCTTGAACCTTACACAGCAAGcCTTACGTCT	3692
QY	1221	AlaAlaValProValLeuAlaLysThrGluValGlyGluGluGlyGluValAspTrpLeu	1240

Db 3593 GCACCGTGGCGGTTCTGGCAAGACTGAGTGGTCAAGAGGGTGAGGTGAGTGGT 3752
 Qy 1241 AspGlyLeuLysValLysGluGlnGluValPheValHisSerGlyProAsnSerGln 1260
 Db 3753 GATGGAGAAAAGTCAAGAAGAACACAGAGGTGTTGTACACTCTGGACCCACAGTCAA 3812
 Qy 1261 LysAlaAlaSerValThrTyrAspSerGluValMetGlyValAlaGlyCysGlnGluLys 1280
 Db 3813 AAGGCTGCTGATGTGACATATGACAGTGAAGTATGATGGAGTGGCGGTTGTCAGGAAAAG 3872
 Qy 1281 GluSerThrGluValGlnSerLeuSerLeuGluGluMetGluThrAspValGlu 1300
 Db 3873 GAGAGTACTGAAGTGCAGAGTCTTAGCTGAGGAGGAGATGGAACTGAGTTGAA 3932
 Qy 1301 LysGluLysArgGluThrLysProGluGlnValSerGluGluGlnGluThrAla 1320
 Db 3933 AAGGAGAAAAGGAGACAAAGCCAGACCAAGTGAAGAGGTGAGCAGGAAACAGCC 3992
 Qy 1321 AlaProGluHisGluGlyThrTyrGlyLysProValLeuThrLeuAspMetProSerSer 1340
 Db 3993 GCTCCTGAGCATGAAGAACTACGGGAAGCCAGTCTGTGACACTTGACATGCCAGCTCA 4052
 Qy 1341 GluArgGlyLysAlaLeuGlySerLeuGlyGlySerProSerLeuProAspGlnAspLys 1360
 Db 4053 GAGAGGGGGAAGGACCTGGGAAGCCTTGGGAAGCCCTTCTCCAGACCAAGACAAA 4112
 Qy 1361 AlaGlyCysIleGluValGlnValGlnSerLeuAspThrThrValThrGlnThrAlaGlu 1380
 Db 4113 GCAGGTGTCATAGAGTTCAGTTCAAGCTCAAGCCCTGGACACACAGTCACTCAACACGACGAA 4172
 Qy 1381 AlaValGluLysValIleGluThrValValIleSerGluThrGlyGluSerProGluCys 1400
 Db 4173 GCTGTGAAAAGGTCATAGAAAACGTTGTGATTTTCAGACAGCAGGTGAAGTCCAGAGTGT 4232
 Qy 1401 ValGlyAlaHisLeuLeuProAlaGluLysSerSerAlaThrGlyGlyHisTrpThrLeu 1420
 Db 4233 GTAGTGCACACTATTATACAGCTGAGAAGTCTCTGCAACGGGTGGCCATGGACTCTT 4292
 Qy 1421 GlnHisAlaGluAspThrValProLeuGlyProGluSerGlnAlaGluSerIleProIle 1440
 Db 4293 CAGCATCAGAGGACAGCGTACCCCTGGGCTGAGTCTCAGGACAGAAATCATCCCAATC 4352
 Qy 1441 IleValThrProAlaProGluSerThrLeuHisProAspLeuGlnGlyGluIleSerAla 1460
 Db 4353 ATAGTAACCTCTGCTCTGCTGAAAGCACCTCATCTGACCTTACAAGGAGAAATAAGCGCA 4412
 Qy 1461 SerGlnArgGluArgSerGluGluAspLysProAspAlaGlyProAspAlaAspGly 1480
 Db 4413 TCCAGAGAGCGCATCAGAGGAGAGGACCAAGCCAGATGCTGCTGCTGATGCTGACGGC 4472
 Qy 1481 LysGluSerThrAlaIleGluLysValLeuLysAlaGluProGluIleLeuGluLeuGlu 1500
 Db 4473 AAGGAGATGACAGCAATCGAAAAGTCTCTCAGGCTGACCTGAGATCTTGAACTTGAG 4532
 Qy 1501 SerLysSerAsnLysIleValLeuAsnValIleGlnThrAlaValAspGlnPheAlaArg 1520
 Db 4533 AGTAAGACCAACAAGATTGCTGAACGTCATTGACAGCAGCGTTGACCAAGTTCGACCGT 4592
 Qy 1521 ThrGluThrAlaProGluThrHisAlaTyrAspSerGlnThrGlnValProAlaCysArg 1540
 Db 4593 ACAGAAACAGCCCGCAAACTCATGCTTATGATTACAGACCCAGGTTCTGCAATGACAG 4652
 Qy 1541 LeuAspSerArgGluProAsnArgCysTrpThrLysMetLysAspAlaLysMetLysHis 1560
 Db 4653 GCTGACAGCAGGAGCCCAACAGATGCTGGACAAAATGAAGTTGCCAAGATGAACAC 4712
 Qy 1561 ProValProGlnProArgGluAspLeuGlnValLeuThrValLeuGluAlaTrpAlaGln 1580
 Db 4713 CCAGTGGCGCAGCCACAGAGGACTTGCAGTCTGACCGTCTGAGGAGCATGGGCTCAG 4772
 Qy 1581 ProArgLysCysLeuProArgLeuGlnLeuLysAlaProValSerLys 1596

Db 4773 CCTCGGAATGCTTGGCGCGCTTGCAGTTGAAGAGCCCGGTGTCAAAG 4820
 RESULT 7
 US-60-360-207-12210
 ; Sequence 12210, Application US/60360207
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig
 ; TITLE OF INVENTION: HUMAN GENOME DISCOVERY SYSTEM AND USES THEREOF
 ; FILE REFERENCE: CLO01321
 ; CURRENT APPLICATION NUMBER: US/60/360,207
 ; CURRENT FILING DATE: 2002-03-01
 ; NUMBER OF SEQ ID NOS: 47235
 ; SEQ ID NO 12210
 ; LENGTH: 6199
 ; TYPE: DNA
 ; ORGANISM: HUMAN
 ; US-60-360-207-12210
 Alignment Scores:
 Pred. No.: 0 Length: 6199
 Score: 6663.00 Matches: 1357
 Percent Similarity: 88.90% Conservative: 76
 Best Local Similarity: 84.18% Mismatches: 156
 Query Match: 82.53% Indels: 23
 DB: 80 Gaps: 11
 US-09-902-432-4 (1-1596) x US-60-360-207-12210 (1-6199)
 Qy 1 MetGlyValAlaGlySerThrGluGlnArgSerProGluGlnProAlaGlySerAspThr 20
 Db 142 ATGGGTGTCAGGAGTCTCCACCGAGCAGCGGAGCCCGGAGCGGAGAGCAGCAGC 201
 Qy 21 ProSerGluLeuValLeuSerGlyHisGlyProAlaAlaGluAlaSerGlyAlaAlaGly 40
 Db 202 CCAGCGAGCTGGAGCTCAGTGGCCATGGCCCGCAGCGGAGGAAAGCCGAGAGTGGCA 261
 Qy 41 AspProAlaAspAlaAspProAlaThrLysLeuProGlnLysAsnGlyGlnLeuSerSer 60
 Db 262 GATCCGCTGACCGGACCCCGCCACAGCTCCACAGAGAAATGTCAGTGTCTGCC 321
 Qy 61 ValAsnGlyValAlaGluGlnGlyAspValHisValGlnGluGlnGlnGluGlyGln 80
 Db 322 GTCAATGGTGTAGCTGAACAAGAGATGCCAGCTCAGAGGAAAGCCAGGATGGGCAA 381
 Qy 81 GluGluGluValValAspGluAspValGlyGlnArgGluSerGluAspValArgGluLys 100
 Db 382 GAGGAAGAGTCACTGTTGAGATGTTGACAGAGAGTCAAGAATGTGAAGAAGAAA 441
 Qy 101 AspArgValGluGluMetAlaAlaAsnSerThrAlaValGluAspIleThrLysAspGly 120
 Db 442 GACCGAGCTAAAGAAATGGCAGCCAGTTCACAGATTTGTTGAAGATATCAAAAGGACGAG 501
 Qy 121 GlnGluGluThrSerGluIleIleGluGlnIleProAlaSerGluAsnValGluGlu 140
 Db 502 CAGGAGAAACACCGGAAATATCGAACACAGATCCCTGCTTCAGAGAGCAATGTGGAGAA 561
 Qy 141 MetValGlnProAlaGluSerGlnAlaAsnAspValGlyPheLysLysValPheLysPhe 160
 Db 562 ATGGCGCAGGCTGCTGAGTCCCAGCTAATGAGTTCAGCTGGCTTCAAGAAAGTATTAAATTT 621
 Qy 161 ValGlyPheLysPheThrValLysLysAspLysAsnGluLysSerAspThrValGlnLeu 180
 Db 622 GTTGGTTTTAAATTCACGGTGAAGAGGATAAAAACGAAAGTCAGATACCGTCCAGCTA 681
 Qy 181 LeuThrValLysLysAspGluGlyGluGlyAlaGluAlaSerValGlyAlaGlyAspHis 200
 Db 682 CTCACATGTCGAAGAGATGAAGCGGAGGAGAGCCCTCCGTCGAGAGCAGGAGACAC 741
 Qy 201 GlnGluProSerValGluThrAlaValGlyGluSerAlaSerLysGluSerGluLeuLys 220
 Db 742 CAAGAGCCCGGAGTGAAGACC---GTCCGCGCAATCAGATCCAAAGAAAGTGAAGTGAAG 798
 Qy 221 GlnSerThrGluLysGlnGluGlyThrLeuLysGlnGlnSerSerThrGluIlePro 240

Db	799	CAATTCACAGAAAGCAAGAGGCACTCTGACCAAGCAACAAGACACAGAAATTTCC	858
Qy	241	LeuGlnAaGluSerAaspGlnAaAaGluGluGlnAaAaLysAaspGluGluGluLys	260
Db	859	CTTCAAGCCGAATTTGGTCAAGGAGCCGAGGAAGAAAGACCAAGATGGAGAAAGAAAC	918
Qy	261	GlnGluLysGluProThrLysSerProGluSerProSerSerProValAsnSerGluThr	280
Db	919	CGAAGAAAGAACTTACCAAGCCCTTGAATCTCCACACGACCCGTGTAGCAATGAGCA	978
Qy	281	ThiSerSerPheLysLysPhePheThrHisGluTTPAaGluTTPaPyrLysThrSer	300
Db	979	ACATCTTCTTCAAGAAATTTCTACACAGGCTGGGCGGCTGGGGCAAGAGCAAC	1038
Qy	301	PheLysLysSerLysGluAaspAaspLeuGluThraGluAaLysAaLysGluGlnGluAa	320
Db	1039	TTCAAGAAACCAAGAAAGATGATCTGGAAACCTTCGGAAGAAAGAAAGAGAGAGCT	1098
Qy	321	GluLysValAaspGluGluGluLysGluLysThrGluProAla--SerGluGluGlnGlu	339
Db	1099	GAAGAAAGTACGAGGAAGAAAGGGGAAAGACAGAGCCAGCCCCAGCGAGAGCAGAG	1158
Qy	340	ProAaGluAaspThrAaspGlnAaAaGluSerAlaAaspThrGluLysValGluLeuPro	359
Db	1159	CTGTGAAAGGACAGACACAGCCAGGTTGTACGCCGACTTATGAAAGGTGAGTTGCT	1218
Qy	360	LeuGluAaspGlnValGluAaspLeuGluAaSerSerGluGluLysCysAlaProLeuAaA	379
Db	1219	TTGGAAGACAGAGTCCGGTGAACCTGGAGACTTTGTG--GAAAGTGTCTCTTTGGCA	1275
Qy	380	ThrGluValPheAaspGluLysMetGluAaHisGlnGluValAaGluValHisVal	399
Db	1276	ACGGAAGTGTGATGAGAAACGGAAGCCCAAGAAAGTTGTTGACAGAGGTCCACGTG	1335
Qy	400	SerThrValGluLysThrGluGluGlnGlnGlyGlyGluAaGluGlyGlyVal	419
Db	1336	AGCACCGTGAAGAAAGATGACGAAAGGGCA--GGAGGACAGAGGTGAGAGGGGATGTG	1392
Qy	420	ValValGluGlyThrGlyLysLeuProProGluLysLeuAaGluProGlnGluVal	439
Db	1393	GTGGTGAAGAGATCGGAGAAATCTTGTCCCTCGAAGAACTGGCTGAAGCCAGAGAGTC	1452
Qy	440	ProGlnGluAaGluProAaGluGluLeuMetLysSerArgGluMetCysValSerGly	459
Db	1453	CCCCGGAAGCTGAGCCTGTGTGAGAGACTGATGAAGACCAAGAAAGTATGCGTCTTGCG	1512
Qy	460	GlyAaspHisThrGlnLeuThrAaspLeuSerProGluGluLysThrLeuProLysHisPro	479
Db	1513	GGTGCCTTACTCTCACTGACAGATCTTAAGTCCGGAAGAAAGATGCTATCCCAACACCC	1572
Qy	480	GluGlyLysLeuValSerGluValAaGluMetLysSerSerGlnLysArgLysValGlnGly	499
Db	1573	GAAAGCATTTGACATGAGGTGAGATCTCTCTCTCAAGAGAAATCAAGGTTCAGAGGA	1632
Qy	500	SerProLeuLysLysLeuPheSerSerSerGlyLeuLysLysLeuSerGlyLysLysGln	519
Db	1633	AGTCCCTCGAAGAAAGCTCTTACAGAGTTGGGCTTAAAGAAAGCTCTCCGGGAAGAGCAG	1692
Qy	520	LysGlyLysArg--GlyGlyGlyGlyAaspGluGluProGlyGluLysGlnHisGlnHis	538
Db	1693	AAGGGGAAGAGAGAGAGGGGGGGAGAGAAAGACCCAGAGAAATACCAACACTTCA	1752
Qy	539	ThrGluSerProGluSerAlaAaspGlnGlnLysGlyGluLysSerAlaSerSerProGlu	558
Db	1753	ACCGAGTCCCAAGAGAGTGTGACAGACCAAGAGGAGAGAGCTTGCTCTTCCCTCA	1812
Qy	559	GluProGluGluThrThrCysLeuGluLysGlyProLeuGluValAProGlnAaspGlyGlu	578
Db	1813	GAGCCCAAGAGATCGCGGTCTGTGAAAGAGGGCCATCGGAAGACCCCAAGGAACGGAA	1872
Qy	579	AlaGluGluGlyThrThrSerAaspGlyGluLysLysArgGluGlyLysThrProTTPAa	598

[illegible]

Qy	956	ThrAlaThrGluThrSerGluAlaLeuArgThrGluGluValThrGluAlaSerGlyAla	975
Db	3007	ACAGCCGACGAAACACACAGAGCGCTCCGCGCTGAAGAACTTACCGAAGCATCAGGGGCA	3066
Qy	976	GluGluThrThrAspMetValSerAlaValSerGlnLeuThrAspSerProAspThrThr	995
Db	3067	GAGAGACCACACACATGGTGTCTGCAGTTTCCCAGCTGTCCGACTCCCGGACACCA	3126
Qy	996	GluGluAlaThrProValGlnGluValGluSerGlyValLeuAspThrGluGluGluGlu	1015
Db	3127	GAGGAGCCACCCAGTTCAGGAGGTAGAGGCTGGCATGCTAGATACGGAAGACAGGAG	3186
Qy	1016	ArgGlnThrGlnAlaIleuGlnAlaValAlaAspLysValLysGluGluSerGlnVal	1035
Db	3187	CGCCAGACGAGGCGCTCTCCAAAGCGGTGCGACAAAGTGAAAGAGCACTCCAGGTG	3246
Qy	1036	ProAlaThrGlnThrValGlnArgThrGlySerLysAlaLeuGluLysValGluGluVal	1055
Db	3247	CCTGCAACCCAGACTCTCGCAGAGAGCAGGCGCGAAGACACTGGAGAAGTGGAGAGGTA	3306
Qy	1056	GluGluAspSerGluValLeuAlaSerGluLysGluLysAspValMetProLysGlyPro	1075
Db	3307	GAGGAGACTCCGAGGTGTGCTACCGAGAAAGAGAGGATGTGTGCGGAGAGACCC	3366
Qy	1076	ValGlnGluAlaGlyAlaGluHisLeuAlaGlnGlySerGluThrGlyGlnAlaThrPro	1095
Db	3367	GTGCAGGAAGCTGAAACTGAGCATCTTGACAGGCTCCGAGACTGTACAGGCTACCCCA	3426
Qy	1096	GluSerLeuGluValProGluValThrAlaAspValAspHisValAlaThrCysGlnVal	1115
Db	3427	GAGAGCCTTGAAGTTCTGAAGTTCACAGAGGATGTAGACCGTGCCACACATGCGAGTT	3486
Qy	1116	IleLysLeuGlnLeuMetGluGlnAlaValAlaProGluSerSerGluThrLeuThr	1135
Db	3487	ATCAAGCACCGACAGCTGATGGAAACAGGCTGTGGCCCTGAGTCATCTGAAACCTTGACA	3546
Qy	1136	AspSerGluThrAsnGlySerThrProLeuAlaAspSerAspThrAlaAspGlyThrGln	1155
Db	3547	GACAGTGAGACAAATGGAAGTACTCCCTCGCAGATTTCAGACACTCCAAACCGGCACACAG	3606
Qy	1156	GlnAspGluThrIleAspSerGlnAspSerLysAlaThrAlaAlaValArgGlnSerGln	1175
Db	3607	CAAGACGAGCCGTTCAGACCCAGGACAGTAATGGCATTTGGCCGTTCAGACAGTTCACAG	3666
Qy	1176	ValThrGluGluGluAlaLaThrAlaGlnLysGluGluProSerThrLeuProAsnAsn	1195
Db	3667	GTCACTGAAGAGAGGAGCGAGCTGCTGTCTCAGACGGAGGGCGCTTCAACA---CCA	3723
Qy	1196	ValProAlaGlnGluGluHisGlyGluGluProGlyArgAspValLeuGluProThrGln	1215
Db	3724	TTTTCCACCCAGGAGAAACACACAGGAAAAAACCCAGGAGGGATGTTCTAGAACCCACA	3780
Qy	1216	GlnGluLeuThrAlaAlaValProValLeuAlaLysThrGluValGlyGlnGluGly	1235
Db	3781	CAAGCGTGGCTGCCGGGCGAGTGCCTATTCTGGCAAGGCTGAGGTGGTCAAGAGGGT	3840
Qy	1236	GluValAspTrpLeuAspGlyGluLysValLysGlu-----GluGlnGlu	1250
Db	3841	GAGGCTGCCAGTTTATGGAGAAAAAGTCAAGAGCGGACAGTGTGTTAAAGAACTCGAG	3900
Qy	1251	ValPheValHisSerGlyProAsnSerGlnLysAlaAlaAspValThrTyAspSerGlu	1270
Db	3901	GTGCCTGTGCACACTGGACCCACAGTCAANAAGACTGTGACTTGACACGTGACAGTGAA	3960
Qy	1271	ValMetGlyValAlaGlyCysGlnGluLysGluSerThrGluValGlnSerLeuSerLeu	1290
Db	3961	GTAATGAAATGGCCAGATGTCAGGAAACTGAGAGTAATGAAGAACAGAGTATTAGCCCG	4020
Qy	1291	GluGluGlyGluMetGluThrAspValGluLysGluLysArgGluThrLysProGluGln	1310
Db	4021	GAGAAAGAGAGATGGGAACCGGACGTTGAAAGAGGAGGAAACAGAGACCAAGACAGACAA	4080

Qy	1311	ValSerGluGluGlyGluGlnGlnThrAlaAlaProGluHiHisGluGlyThrTyRgLyLys	1330
Db	4081	GCCAGTCAAGAACATCAGCAGGAGAAACAGCTGCTCTCTGACGATGAAGGAACCCACCCCTAAG	4140
Qy	1331	ProValLeuThrLeuAspMetProSerSerGluArgGlyLysAlaLeuGlySerLeuGly	1350
Db	4141	CCAGTCTGTACAGCTGACATGCTCTACTCAGAGAGGGGAAGGACATGGCGAGCCTTGAA	4200
Qy	1351	GlySerProSerLeuProAspGlnAspLysAlaGlyCysIleGluValGlnValGlnSer	1370
Db	4201	GGAAAGCCCTTCTCTCCACAGCAACAGAAACAGATTGCTAGAGGTTTCAAGTTCAAAGC	4260
Qy	1371	LeuAspThrThrValThrGlnThrAlaGluAlaValGluLysValIleGluThrValVal	1390
Db	4261	TCAGACACACAGTCACTCAAAACAACCGAAGCTGTGAAAAAGGTGCAAGAAACCTGTGGCA	4320
Qy	1391	IleSerGluThrGlyGluSerProGluCysValGlyAlaHisLeuLeuProAlaGluLys	1410
Db	4321	ACTTCAGAGATGATGAAGATTTGGAGTGTGCAAGTGGCGCAATCATTTACCAAGCTGAGAAG	4380
Qy	1411	SerSerAlaThrGlyGlyHisTrpThrLeuGlnHisAlaGluAspThrValProLeuGly	1430
Db	4381	CTCTCCGAAACCGTGGCTACGGGACTCTTCAGCATGGAGAGGACCGCTGCCCCAGGGG	4440
Qy	1431	ProGluSerGlnAlaGluSerIleProIleIleValThrProAlaProGluSerThrLeu	1450
Db	4441	CCTGAGTCTCAGGCAGAGTCCATCCCAATAATAGTAACCTCTGCTCTCTGAAAGCATCCTA	4500
Qy	1451	HisProAspLeuGlnGlyLurIleSerAlaSerGlnArgGluArgSerGluGluGluAsp	1470
Db	4501	CATTCTCAGCTTCAAGAGAAGTGAAGCGCATCCCAAGAAACAGACATCAGATGAAGATAAC	4560
Qy	1471	LysProAspAlaGlyProAspAlaAspGlyLysGluSerThrAlaIleGluLysValLeu	1490
Db	4561	AAGCCAGATGCTGGTCTGATGCTCCGCCAGAGAGAGTGTCAGCAAGAGAGAAAATCCTC	4620
Qy	1491	LysAlaGluProGluIleLeuGluLeuGluSerLysSerAsnLysIleValLeuAsnVal	1510
Db	4621	AGGGCTGAACCTCAGATCTTTGGAACCTTGAGAGTAAGAGACAATAGATTGTCCAGAGTGC	4680
Qy	1511	IleGlnThrAlaValAspGlnPheAlaArgThrGluThrAlaProGluThrHisAlaTyr	1530
Db	4681	ATCCAGACAGCCGTCACACAGTTTGACGCTGACAGGCTGACAGCCAGGAGCAACACAGATGCTG	4740
Qy	1531	AspSerGlnThrGlnValProAla-CysArgLeuAspSerArgGluProAsnArgCysTr	1550
Db	4741	GATTTACAGAAATCAGGTCTCTGTGATGAGGCTGACAGCCAGGAGGACACACAGATGCTG	4800
Qy	1550	pThrLysMetLys-----AspAlaLysMetLysHisProValProGlnPr	1565
Db	4801	GACAAAGATGAAGCGACCTTCAAGTCTCCCCCAAGATGAACACATCAGTGCCGTAGCC	4860
Qy	1565	oArgGluAspLeuGlnValLeuThrValLeuGluAlaTrpAlaGlnProArgLysCysLe	1585
Db	4861	CAGGAAGGACTTGCGGTTTCTGTATGTTCTGAAGGCATGCAAGAGGCTTCAAGAAATGATC	4920
Qy	1585	uProArgLeuGlnLeuLysAlaProValSerLys	1596
Db	4921	ACCAAGCTTGCAGTTGAAAGTGCAGTGTCAAAG	4954
RESULT 8			
US-08-978-277A-1			
; Sequence 1, Application: US/08978277A			
; GENERAL INFORMATION:			
; APPLICANT: Gelman, Irwin H.			
; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE			
; NUMBER OF SEQUENCES: 20			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond			
; STREET: 30 Rockefeller Plaza			
; CITY: New York			
; STATE: NY			
; COUNTRY: USA			

```

ZIP: 10112-0228
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/978,277A
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/665,401
FILING DATE: 18-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Richard S
REGISTRATION NUMBER: 26,154
REFERENCE/DOCKET NUMBER: A30558 - 165/34008
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-408-2558
TELEFAX: 212-765-2519
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5074 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHEetical: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
US-08-978-277A-1

Alignment Scores:
Pred. No.: 6,1e+300 Length: 5074
Score: 6005.00 Matches: 1232
Percent Similarity: 97.35% Conservative: 7
Best Local Similarity: 96.78% Mismatches: 27
Query Match: 74.38% Indels: 14
DB: 13 Gaps: 4

US-09-902-432-4 (1-1596) x US-08-978-277A-1 (1-5074)
QY 339 GtutSThgluproAlaSerGtugIngluproAlaGluaspThraSpGlnaIaarg 348
DB 2 GAAAAGACGAGCCGCTCGAGGAGGAGCGGCGAGGAGACACAGCCAGCCAGG 61
QY 349 LeuSerAlaSPtYrGtUlySValGtUleuproLeuGluaspGlnaIaGlyaspLeuGlu 368
DB 62 TTGTGAGGACTACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 121
QY 369 AlaSerSerGluGluuycysAlaProlaAlaThrGluValPheaspGluysMetGlu 388
DB 122 GATGCTCGAGGAGAGAGTGTGCTCTTTGGACACGAAAGTGTTAATGAGAAAGATGAA 181
QY 389 AlaHseGlnGluValaIaGluValaIaHsaIserThrValGluYlThrGtugInglu 408
DB 182 GCCACCAAGAGTTGTTCCAGAGGTCACAGTGAACCCGTGAGAAAGACAGAGAGAG 241
QY 409 GlnGlyGlyGlyGlyGluIaGluGlyGlyValaIaGluGlyThrGlyGlyIuSerLeu 428
DB 242 CAGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 301
QY 429 ProProGluLysLeuAlaGluProGlnGluValaProGlnGluIaGluProAlaGluGlu 448
DB 302 CCCCTCGAAGAACTGCTGAGAGCCCGAGAGGTCCTCCAGGAACCTAGCTGAGAGAG 361
QY 449 LeuMetLysSerArgGluMetCysValIserGlyGlyAspHisThrGlnLeuThrAspLeu 468
DB 362 CTGATGCAAGACGAGAGAGATGTGTCTCTGAGGAGACCAACACTGAAGAGAGCTTA 421
QY 469 SerProGluGluYlThrLeuProLysHisProGluGlyIleValIserGluValaGluMet 488

```

```

DB 422 AGTCTGAAGAGAGACCGCTGCCAAACACCCAGAGGACTTGTGAGAGTGAATG 481
QY 489 LeuSerSerGlnGluValaGlyIleValaGlnGlySerProLeuLysLeuPheSerSer 508
DB 482 CTGTCTCTCAGAAAGAAATCAAGTACAGGAAAGTCCCTTGAAGAACTCTTCAGTAGC 541
QY 509 SerGlyLeuLysLeuSerGlyYlYlsGlnGlySerGlyValaGlyGlyGlyGlyasp 528
DB 542 TCAGGCTTAAAGAAAGCTGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 601
QY 529 GluGluProGlyGluYlThrGlnHisIleHisThrGluSerProGluSerAlaSpGluGln 548
DB 602 GAAGAGCTTGAGAAATACCAACATTCACACCAATCCAGAGAGAGCTGATGAGAGAG 661
QY 549 LysGlyGluSerSerAlaSerSerProGluGluProGluGluThrCysLeuGluLys 568
DB 662 AAGGAGAGAGAGCTGCTGCTGCTGCCAGAGAGCTGAGAGAGACCACTGTGAGAGAA 721
QY 569 GlyProLeuGluIaIaProGlnaSpGlyGluIaGluGlyThrThrSerAspGlyGlu 588
DB 722 GGGCCGCTGAGAGACCCAG-GATGGGAGAGCTGAGAGAGAGAACTACTTGT--GAGAG 778
QY 589 LysLysArgGluGlyIleThrProThrAlaSerPheLysLysMetValThrProLysLys 608
DB 779 AAGAGAGAGAGAG--ATCACTCTCTGGGCACTCTTCAAAAGATGTGTACACCCAGAGAA 836
QY 609 ArgValaArgArgProSerGluSerAspLysGluGluGluGluValaLysSerAla 628
DB 837 CGGTC-CGAAAGCTTCTGAGAGTGACAAAGAGAGAGAGCTGAGAGAGCTCAAGAGCCGC 895
QY 629 ThrLeuSerSerThrAspSerThrValSerGluMetGlnaSpGluValaLysThrValaGly 648
DB 896 ACCTTGCTCTCCAGATGACAGATGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 955
QY 649 GluGluGlnLysProGluGluProLysArgValaAspThrSerValSerTrpGluIa 668
DB 956 GAGGAAACAAAGCCAGAGAGAACCAAGGCTGAGTGTGATCTTCACTGTCTTGGAGAGCA 1015
QY 669 LeuIleCysValaGlySerSerLysLysArgLaaArgLysAlaSerSerSerAspSpGlu 688
DB 1016 CTGATTGTGTGCGATCATCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1075
QY 689 GlyGlyProArgThrLeuGly-GlyAspSerHisArgAlaGluGluIaAspLysAspLys 708
DB 1076 ---GGGCAAGAGACACTGGAGGAGGAGACATCAAGAGAGAGGAGGAGGAGAGAGCA 1132
QY 708 sGluIaGlyThrAspAlaValaProAlaSerThrGlnGluGlnaSpGlnaIaGlnGlySe 728
DB 1133 AGAAAGCCGA-ACAGAGGCTGTCTTCCAGACCCAGAGAGAGAGAGAGAGAGAGAGAG 1191
QY 728 rSerSerProGluProAlaGlySerProSerGluGluGlyGluValaIserThrTrpGluSe 748
DB 1192 TTCTCACCCGAGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1251
QY 748 rPheLysArgLeuValThrProArgLysLysSerLysSerLysLeuGluGluVala-- 767
DB 1252 ATTTAAAGATTATGCTCACTCCAAAGAAAAATCCAAAGTCAAAATCGGAGAGAGAA--AG 1310
QY 768 ---GluAspSerSerVal---GluGlnLeuSerThrGluIleGluProSerArgGluGlu 785
DB 1311 CCGAAGAGACTTACTGTAGAGAGAGGTTGTCCCTAGATCGAAACCGTGTAGAGAGAA 1370
QY 785 uSerTrpValSerIleLysLysPheIleProGlyArgArgLysLysArgAlaAspGlyLys 805
DB 1371 ATCTTGGGTTTCATTAGAAATTCATCCCGGAGAGGAGAGAGAGAGAGAGAGAGAG 1430
QY 805 s-GlnGluGlnAlaThrValaGluAspSerGlyProValaGluIleAsnGluAspAspPro 825
DB 1431 GGCAAGAACAGAGCCACTGTGAGAGAGCTCAGAGGAGTGAAGATTAATGAGAGAGAGCTG 1490
QY 825 snValProAlaValaIaProLeuSerGluYlThrAsnAlaValaGluArgGluLysMetGlu 845

```

Db 1491 ATGTCACGCGCTGCTGCTCTGCTGAGTATGATGTCAGTGGAGGAGGAGATGGAAG 1550
Qy 845 laGlnGlyAsnThrGluLeuProGlnLeuLeuGlyAlaValTyrValSerGluGluLeuS 865
Db 1551 CCCAGGGGAATGCGGAGCTGCCAGCTGCTGGGGCTGTGTA--GTGTCGAGGAGCTCA 1607
Qy 865 erLysThrLeuValHisThrValSerValAlaValIleAspGlyThrArgAlaValThrS 895
Db 1608 GTAAGACTCTGGTCCACACTGTGAGTGTGCGAGTCATGTGATGGGACCAGGGCAGTCACCA 1667
Qy 885 erValGluGluArgSerProSerTriPileSerAlaSerValThrGluProLeuGluHist 905
Db 1668 GTGTCGAAGAGCGGTCTCTCTGTGGATATCCGCTTCGTAACGAACCTCTTGAACACA 1727
Qy 905 hrAlaGlyGluAlaMetProProValGluGluValThrGluLysAspIleAlaGluG 925
Db 1728 CAGCGGAGAGGACCATGCCACCTGTTCAAGAGGTCACTGAAAAAGACATCATTTGCAGAAG 1787
Qy 925 luThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHisAspMetValT 945
Db 1788 AAACCTCTGTGCTCACCCAGACGTTACAGAGGGTAAGATGCCCATGACGATGGTCA 1847
Qy 945 hrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeuA 965
Db 1848 CCAGTGAAGTGGATTTACCTCAGAACCTGTGACAGCCACAGAGACCTCAGAGGCTCTCC 1907
Qy 965 rgThrGluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSerAlaV 985
Db 1908 GTACTGAAGAAGTTACCGAAGCATCGGGGCGGAAGAGACCACAGACATGGTGTCCGACG 1967
Qy 985 alSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGluValG 1005
Db 1968 TTTCCCAAGCTGACTGACTCCCGACACACCAGAGGAAGCCACCCCAAGTTCCAGAGGTAG 2027
Qy 1005 luSerGlyValLeuAspThrGluGluGluArgGlnThrGlnAlaIleLeuGlnAlaV 1025
Db 2028 AGGGTGTGTGCTAGATACAGAGAGAGGAGGCCAGAGCCAGGCCATCCTCAAGCCG 2087
Qy 1025 alAlaAspLysValLysGluSerGlnValProAlaThrGlnThrValGlnArgThrG 1045
Db 2088 TTGCAGACAAGGTGAAAGAGGAGTCCCGAGGTGCGCTGCTCAACCCAGACTGTGCAGAGAACG 2147
Qy 1045 lySerLysAlaLeuGluLysValGluGluValGluGluAspSerGluValLeuAlaSerG 1065
Db 2148 GGTCAAAAGCACTGGAGAAGCTTTAGAGAGGTAGAGGAGGACTCCGAAGTGTGCGCTTCGG 2207
Qy 1065 luLysGluLysAspValMetProLysGlyProValGlnGluAlaGlyValAlaGluHisLeuA 1085
Db 2208 AGAAAGAGAAGGAGCTTATGCCGAAAGACCCGTCGAGGAAGCTGGAGCTGAGCATCTTG 2267
Qy 1085 laGlnGlySerGluThrGlyGlnAlaThrProGluSerLeuGluValProGluValThrA 1105
Db 2268 CACAGGCTCTGAGACTGGACAGGCTACTCCAGAGAGCCCTTGAAGTTCTCTGAAGTCAACG 2327
Qy 1105 laAspValAspHisAlaThrCysGlnValIleLysLeuGlnGlnLeuMetGluGlnA 1125
Db 2328 CAGATGTAGACCATGTGCGCACGTCGCCAGGTATCAAGCTCCAGCAGCTCATGGAACAGG 2387
Qy 1125 laValAlaProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrProL 1145
Db 2388 CCGTGGCCCTTGAGTCTATCCGAACCTTGACAGACAGTGGACAAATGGGAAGCACTCCCT 2447
Qy 1145 euAlaAspSerAspThrAlaAspGlyThrGlnGlnAspGluThrIleAspSerGlnAsps 1165
Db 2448 TAGCAGATTACAGACACTGCAGATGGGACACAGCAAGATGAACCATTTGACAGCCAGGACA 2507
Qy 1165 erLysAlaThrAlaAlaValArgGlnSerGlnValThrGluGluGluAlaAlaThrAlaG 1185
Db 2508 GTAAAGCCACTGACGCTGTGAGGAGTCAAGGTCACAGAGGAGAGGCGGCTACTGCTC 2567
Qy 1185 lnLysGluGluProSerThrLeuProAsnAsnValProAlaGlnGluGluHisGlyGluG 1205
Db 2568 AGAAAGAGGAGCCCTTCACACTACCTTAATAATGTTCCAGGCCAGGAAAGACATGGGGAGG 2627

Qy 1205 luProGlyArgAspValLeuLeuProThrGlnGlnGluLeuThrAlaAlaValProV 1225
Db 2628 AACCAAGAAAGAGATGTTCTTGAACCTACACAGCAAGAGCTTGTCTGCGACCGCTCCCG 2687
Qy 1225 alLeuAlaLysThrGluValGlyGlnGluGlyValAspTrpLeuAspGlyGluLysV 1245
Db 2688 TCTGGCAAAAGACTGAGGTGGGTCAAGAGGGTGAGGTGACTGGTTGGATGGGAAAAAG 2747
Qy 1245 alLysGluGlnGlnValPheValHisSerGlyProAsnSerGlnLysAlaAlaAspV 1265
Db 2748 TCAAGAAGAAACAGAGAGGTGTTGTACACTCTGGACCCCAACAGTCAAAAGGCTGCTGATG 2807
Qy 1265 alThrTyrAspSerGluValMetGlyValAlaGlyCysGlnGluLysGluSerThrGluV 1285
Db 2808 TCACATATCACAGTGAAGTGTGGAGTGGCGGTCTCAGGAAAAAGGAGAGTACTCAAG 2867
Qy 1285 alGlnSerLeuSerLeuGluGlyGluMetGluThrAspValGluLysGluLysArgG 1305
Db 2868 TGCAGAGTCTTTAGCCTGGAGGAGGAGATGGAACCTGACGTTGAAAAGGAGAAAAAGG 2927
Qy 1305 luThrLysProGluGlnValSerGluGluGlyGlnGlnGluThrAlaAlaProGluHisG 1325
Db 2928 AGACAAAGCCAGAGCAAGTGAGTGAAGAAGGTGAGCAGGAAACAGCCGCTCCTGAGCATG 2987
Qy 1325 luGlyThrTyrGlyLysProValLeuThrLeuAspMetProSerSerGluArgGlyLysA 1345
Db 2988 AAAGGAACCTACGGAAGCCAGTCTGACACTTGACATGCTGACAGTCCAGAGAGGGGAAGG 3047
Qy 1345 laLeuGlySerLeuGlyGlySerProSerLeuProAspGlnAspLysAlaGlyCysIleG 1365
Db 3048 CACTGGGAAGCCTTGAGGAAGCCCTTCTCTCCAGACCAAGACAGGAGGTGTCATAG 3107
Qy 1365 luValGlnValGlnSerLeuAspThrThrValThrGlnThrAlaGluAlaValGluLysV 1385
Db 3108 AGGTTCAAGTTCAAGCCTGGACACAACTGACTCAACAGCAGGAAGCTGTGAAAAAGG 3167
Qy 1385 alIleGluThrValValIleSerGluThrGlyLysSerProGluCysValGlyAlaHisL 1405
Db 3168 TCATGAACCGGTGTGATTTTCAGACAGAGTGAAGTCCAGAGTGTGTAGGTGCACACT 3227
Qy 1405 euLeuProAlaGluLysSerSerAlaThrGlyGlyHisThrThrLeuGlnHisAlaGluA 1425
Db 3228 TATTCACAGCTGAGAAGTCTCTGCAACGGGTGGCCACTGGACTCTTCAGCATGCGAGAG 3287
Qy 1425 spThrValProLeuGlyProGluSerGlnAlaGluSerIleProIleIleValThrProA 1445
Db 3288 ACACGGTACCCCTGGGCGCTGAGTCTCAGGCAGATCCATCCCAATCATAGTAACCTCCTG 3347
Qy 1445 laProGluSerThrLeuHisProAspLeuGlnGlyGluIleSerAlaSerGlnArgGluA 1465
Db 3348 CTCCTGAAAGCACCCCTACATCTCTGACCTCAAGGAGAAATNAGCGCATCCACAGAGAGC 3407
Qy 1465 rgSerGluGluGluAspLysProAspAlaGlyProAspAlaAspGlyLysGluSerThrA 1485
Db 3408 GATTCAGAGGAAGAGGACAAAGCCAGATGCTGGTCTGTGCTGACGCAAGGAGAGTACAG 3467
Qy 1485 laIleGluLysValLeuLysAlaGluProGluIleLeuGluLeuGluSerLysSerAsnL 1505
Db 3468 CAATCCAAAGTCTTCAAGGCTGAACCTGAGATCTCTGGAACCTTGAGATTAAGAGCAACA 3527
Qy 1505 ysIleValLeuAsnValIleGlnThrAlaValAspGlnPheAlaArgThrGluThrAlaP 1525
Db 3528 AGATTGTGCTGAACGTCATTTCAGACAGCGCTTGACCACTCGCACGTACAGAAACAGCCC 3587
Qy 1525 roGluThrHisAlaTyrAspSerGlnThrGlnValProAlaCysArgLeuAspSerArgG 1545
Db 3588 CCGAAACTCATGCTTATGATTCACAGACCAGGTTCTCTGCAATGCGCTTGGACAGCAGGG 3647
Qy 1545 luProAsnArgCysTyrThrLysMetLysAspAlaLysMetLysHisProValProGlnP 1565
Db 3648 AGCCCAACAGATGCTCGACAAAAAATGAAAGTTGCCAAGATGAAACACCCAGTCCCGCAGC 3707

Db 1192 TTCTCACCCGAGCCAGCGGAGCCCTTCCGAAGGGGAAGGTGTCTCCACTTGGGAGTC 1251
Qy 748 rPheLysArgLeuValThrProArgLysLysSerLysLeuGluGluLysAla-- 767
Db 1252 ATTTAAAGATTAGTCACTCCAGAAAAAATCCAAAGTCAAACTGGAAGAGAAGA-AG 1310
Qy 768 ----GluAspSerVal---GluGlnLeuSerThrGluIleGluProSerArgGluGI 795
Db 1311 CCGGAAGGACTCTAGTGTGTAGGAGCAGGTTGTCCACTGAGATCGAACCCGTTAGAGAAGA 1370
Qy 785 uSerTrpValSerIleLysLysPheIleProGlyValArgLysLysArgAlaAspGlyLY 805
Db 1371 ATCTTGGGTTTCCATTAGAAATTCATCCCGGAGCGGCGGAAGAAAGGCGAGATGGGAA 1430
Qy 805 s-GlnGluGlnAlaThrValGluAspSerGlyProValGluIleAsnGluAspProA 825
Db 1431 GGCAAGAAACAAGCACTGTGGAAGACTCAGGGCCAGTGGAGATAAATGAGACGAGCCTG 1490
Qy 825 snValProAlaValValProLeuSerGluTyrAsnAlaValGluArgGluLysMetGluA 845
Db 1491 ATGTCCAGCGCGTGGTCTCTGTCTGAGTATGATGCAGTGGAGGAGGAAGATGGAAG 1550
Qy 845 laGlnGlyAsnThrGluLeuProGlnLeuLeuGlyAlaValTyrValSerGluGluLeus 865
Db 1551 CCAGGGGAATGCGGAGCTGCCAGCTGCTGGGGCTGTGTA---GTGTCCGAGAGCTCA 1607
Qy 865 erLysThrLeuValHisThrValSerValAlaValIleAspGlyThrArgAlaValThrS 885
Db 1608 GTAAGACTCTGGTCCACACTGTGAGTGTGTCAGTTCATTGATGGGAGCCAGGCGAGTCA 1667
Qy 885 erValGluGluArgSerProSerTrpIleSerAlaSerValThrGluProLeuGluHisT 905
Db 1668 GTGTGGAAGACGGGTCTCTTCGTGGATATCCGCTTCCGTAACAGAACCTCTTTGAACACA 1727
Qy 905 hrAlaGlyGluAlaMetProProValGluGluValThrGluLysAspIleAlaGluG 925
Db 1728 CACGGGGAAGGCATGCCACCTGTTGAGAGGTCACTGAAAAGAGCATCATTTGCGAAG 1787
Qy 925 luThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHisAspAspMetValT 945
Db 1788 AAACCTCTGTGCTCACCCAGACGTACCAGAGGTTAAAGATGCCATGACGACATGCTCA 1847
Qy 945 hrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeuA 965
Db 1848 CCAGTGAAGTGGATTTTCACTCAGAAGCTGTGACAGCCACAGACACCTCAGAGGCTCTCC 1907
Qy 965 rgThrGluGluValThrGluAlaSerClyAlaGluGluThrThrAspMetValSerAlav 985
Db 1908 GTACTGAAGAAGTTTACCAGAGCATCGGGGCGGAAGAGACACAGACATGGTGTCCGAG 1967
Qy 985 alSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGluValG 1005
Db 1968 TTTCCAGCTGACTGACTCTCCACAGACACACAGAGGAGGCCACCCCAAGTTCCAGAGGTAG 2027
Qy 1005 luSerGlyValLeuAspThrGluGluGluGluArgGlnThrGlnAlaIleLeuGlnAlav 1025
Db 2028 AGGCTGTGTGTAGATACAGAGAAGAGAGGAGCGCCAGACGACGAGGCTCTCCAGCCG 2087
Qy 1025 alaAlaAspLysValLysGluSerGlnValProAlaThrGlnThrValGlnArgThrG 1045
Db 2088 TTGCAGACAAGGTGAAGAGAGTCCCAGGTGCTGCAACCCAGACTGTCCAGAGAACGG 2147
Qy 1045 lySerLysAlaLeuGluLysValGluGluValGluGluAspSerGluValLeuAlaSerG 1065
Db 2148 GGTCAAAAGCACTGGAGAAGTTGAGAGGTAGAGGAGGACTCCGAAGTGTGCTGTCCG 2207
Qy 1065 luLysGluLysAspValMetProLysGlyProValGlnGluAlaGlyAlaGluHisLeuA 1085
Db 2208 AGAAGAGAAGGACGTTATGCCGAAGGACCCGCTGCAGGAAGCTGGAGCTGAGCATCTTG 2267
Qy 1085 laGlnGlySerGluThrGlyGlnAlaThrProGluSerLeuGluValProGluValThrA 1105

Db 2268 CACAGGGCTCTGAGACTGGACAGGGCTACTCCAGAGAGCCCTTGAAGTTCCTGAAGTCACAG 2327
Qy 1105 laAspValaAspHisValAlaThrCysGlnValIleLysLeuGlnLeuMetGluGlnA 1125
Db 2328 CAGATGTAGACCATGTGCGCACAGTGCAGGTTATCAAGCTCCAGCAGCTGATGTAACAGG 2387
Qy 1125 laValaAlaProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrProL 1145
Db 2388 CCGTGGCCCTCTGAGTCATCCGAACCTTGACAGACAGTGGAGCAAAATGGAAAGCACCTCCT 2447
Qy 1145 euAlaAspSerAspThrAlaAspGlyThrGlnGlnAspGluThrIleAspSerGlnAsps 1165
Db 2448 TAGCAGATTACAGACTGCAGATGGACACAGCAAGATGAACCATTTGACAGCCAGGACA 2507
Qy 1165 erLysAlaThrAlaAlaValaArgGlnSerGlnValThrGluGluGluAlaAlaThrAlaG 1185
Db 2508 GTAAAGCCACTGCAGCTGTGAGCAGTACAGGTTCACAGAAAGAGGCGGTACTGCTC 2567
Qy 1185 lnyLysGluGluProSerThrLeuProAsnAsnValProAlaGlnGluGluHisGlyGluG 1205
Db 2568 AGAAAGAGGAGGCTTCGACACTACCTTAATGTTCCAGCCAGGAGGAACATGGGGGAAG 2627
Qy 1205 luProGlyArgAspValLeuGluProThrGlnGlnGluLeuThrAlaAlaValProV 1225
Db 2628 AACCAGGAAGAGATGTTCTTGAACCTACAGCAAGAGCTTGTCTGTCGACCGCTGCCCG 2687
Qy 1225 albeuAlaLysThrGluValGlyGlnGluGlyValAspTrpLeuAspGlyGluLysV 1245
Db 2688 TCTGGCAAAAGACTGAGGTGGGTCAAGAGGTTGAGTTGACTGTTGGATGGAGAAAAG 2747
Qy 1245 allysGluGlnGlnGluValPheValHisSerGlyProAsnSerGlnLysAlaAlaAspV 1265
Db 2748 TCAAGAAAGAACAGAGGAGTGTGTTGTACACTGGACCCACAGTCAAAAGGCTGCTGATG 2807
Qy 1265 alThrTrpAspSerGluValMetGlyValaGlyCysGlnGluLysGluSerThrGluV 1285
Db 2808 TGCATATGACAGTGAAGTGAAGTGGGAGTGGCGGTTGTTCAGAAAAGGAGAGTACTGAAG 2867
Qy 1285 alGlnSerLeuSerLeuGluGluGlyMetGluThrAspValGluLysGluLysArgG 1305
Db 2868 TGCAGAGTCTTAGCTGGAGGAGGAGAGATGGAAACTGACGTTGAAAAGGAGAAAAGGG 2927
Qy 1305 luThrLysProGluGlnValSerGluGluGlyGlnGlnThrAlaAlaProGluHiG 1325
Db 2928 AGCAAAAGCCAGACAAGTGAAGTGAAGAGGTGAGCAGGAAAACAGCGCTCTCTGAGCATG 2987
Qy 1325 luGlyThrTyrGlyLysProValLeuThrLeuAspMetProSerSerGluAArgGlyLysA 1345
Db 2988 AAGGAACTACGGGAAGCCAGTCTTGACACTTGACATGCCAGCTCAGAGAGGGGGAAGG 3047
Qy 1345 laLeuGlySerLeuGlyGlySerProSerLeuProAspGlnAspLysAlaGlyCysIleG 1365
Db 3048 CACTGGGAAGCCTTGGAGGAGCCCTTCTCTCCAGACCAAGACAGAGGTTGTCATAG 3107
Qy 1365 luValGlnValGlnSerLeuAspThrThrValThrGlnThrAlaGluAlaValGluLysV 1385
Db 3108 AGGTTCAAGTTCAAAGCCTGGACACCAACAGTCACTCAAAACAGCAGAAGCTGTGGAAAAGG 3167
Qy 1385 alIleGluThrValValIleSerGluThrGlyGluSerProGluCysValGlyAlaHisL 1405
Db 3168 TCATAGAAACGGTTGTGATTTTCAGACAGAGGTGAAAAGTCCAGAGTGTGTAGGTGCACACT 3227
Qy 1405 euLeuProAlaGluLysSerSerAlaThrGlyGlyHisTrpThrLeuGlnHisAlaGluA 1425
Db 3228 TATTACAGCTGAGAAGTCTCTGCAACGGGTGGCCACTGGACTCTTCAGCATGCGAGG 3287
Qy 1425 spThrValProLeuGlyProGluSerGlnAlaGluSerIleProIleIleValThrProA 1445
Db 3288 ACACGGTACCCCTGGGCGCTGAGTCTCAGCGCAAAATCCATCCCAATCATAGTAACCTCTG 3347
Qy 1445 laProGluSerThrLeuHisProAspLeuGlnGlyLysLeuSerAlaSerGlnArgGluA 1465
Db 3348 CTCTTGAAGACCCCTTACATCTCTGACCTCAAGGAGAAAATAAGCGCATCCCGAGAGAGC 3407

QY 629 ThrLeuSerSerThrAspSerThrValSerGluMetGlnAspGluValIysThrValGly 648
DB 896 ACCTGTGCTCCCTAGATAGCAGGTGCAGAAATGCAGATGAAGTCAAAACTGTGTGT 955
QY 649 GluGluGlnLysProGluGluProLysArgValAspThrSerValSerTrpGluAla 668
DB 956 GAGGACAAAAGCCAGAGGAACCAAGCGTAGGGTGATCTTCACTGCTTGGGAAGCA 1015
QY 669 LeuIleCysValGlySerSerLysLysArgAlaArgLysAlaSerSerSerAspGlu 688
DB 1016 CTGATTTGTCTCGATCATCCAGAGAGAGCAAGGAAGCATCTCTTCAGATATAAGA 1075
QY 689 GlyGlyProArgThrLeuGly-GlyAspSerHisArgAlaGluAlaSerLysAspLys 708
DB 1076 ---GGGCCAAGGACACTGGGAGGGGGACGTCAAGAGCAGAGAGGCCAGCAAGACAA 1132
QY 708 sGluAlaGlyThrAspAlaValProAlaSerThrGlnGluGlnAspGlnAlaGlnGlySe 728
DB 1133 AGAAGCCGA-ACAGACGCTGTTCTTCGCAGCAACCAGGACAGGACCAAGCGCAGGAAG 1191
QY 728 rSerSerProGluProAlaGlySerProSerGluGlyGluGlyValSerThrTrpGluSe 748
DB 1192 TTCCTCACCCGAGCCAGCGGAAGCCCTCCGAAGGGGAAGGTGTCTCCACTTTGGGAGTC 1251
QY 748 rPheLysArgLeuValThrProArgLysLysSerLysSerLysLeuGluGlyLysAla-- 767
DB 1252 ATTTAAAGATTAGTCTCTCCAGAAAGAAAATCCAACTCAAACTGGGAAGAGAAG-AG 1310
QY 768 ---GluAspSerSerVal---GluGlnLeuSerThrGluIleGluProSerArgGluG 785
DB 1311 CCGAAGGACTCTAGTTGTAGGAGCAGGTTGTCTCACTGAGATCGAACCGTGTAGAGAAGA 1370
QY 785 uSerTrpValSerIleLysLysPheIleProGlyArgArgLysLysArgAlaAspGlyLys 805
DB 1371 ATCTTGGGTTTCCATTAAAGAAATTCATCCCGGAGCGCGAAGAAAGGCGCAGATGGAA 1430
QY 805 s-GlnGluGlnAlaThrValGluAspSerGlyProValGluIleAsnGluAspAspProA 825
DB 1431 GGAAGAACAAGCCACTGTGGAGACTCAGGGCCAGTGGAGATAAATAGAGCAGCGCTG 1490
QY 825 snValProAlaValProLeuSerGluTyrAsnAlaValGluArgGluLysMetGluA 845
DB 1491 ATGTCCAGCGCTGCTGCTCTGTCTGAGTATGATGCAGTGGAGGAGAGAAGATGGAA 1550
QY 845 laGlnGlyAsnThrGluLeuProGlnLeuLeuGlyAlaValTyrValSerGluGluLeuS 865
DB 1551 CCAGGGGAATGCGGAGCTGCCAGCTGCTGGGCTGTGTA---GTGTCCGAGGAGCTCA 1607
QY 865 erLysThrLeuValHisThrValSerValAlaValIleAspGlyThrArgAlaValThrS 885
DB 1608 GTAAGACTCTGGTCCACACTGTGAGTGTGCGAGTCAATTGATGGGACCAGGGCGACTCA 1667
QY 885 erValGluGluArgSerProSerTrpIleSerAlaSerValThrGluProLeuGluHisT 905
DB 1668 GTGTGGAAGAGCGGTCTCTTGTGGGATATCCGCTTCCGTAACAGAACCTCTTGAACACA 1727
QY 905 hrAlaGlyGluAlaMetProProValGluGluValThrGluLysAspIleAlaGluG 925
DB 1728 CAGCGGGAAGAGCCATCCACCTGTTGAAGAGGTCACTGAAAGAGACATCATTTGCAGAG 1787
QY 925 luThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHisAspAspMetValT 945
DB 1788 AAATCTGTGTCTCACCCAGACGTTACAGAGGGTAAGATGCCATGACGACATGTGTCA 1847
QY 945 hrSerGluValAlaAppPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeuA 965
DB 1848 CCAGTGAAGTGGATTTCACCTCAAGAGCTGTGACAGCCACAGAGACCTTCAGAGCTCTCC 1907
QY 965 rgThrGluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSerAlav 985
DB 1908 GTACTGAAGAGATTACCGAAGCATCGGGGCGGAGAGACACACAGCATGTGTCCCGAG 1967
QY 985 alSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGluValG 1005

DB 1968 TTTCCAGCTGACTGACTCCCCAGACACACAGAGAGAAAGCCCCAGTTCAGGAGGTAG 2027
QY 1005 luSerGlyValLeuAspThrGluGluGluArgGlnThrGlnAlaIleLeuGlnAlav 1025
DB 2028 AGGGTGGTGTCTAGATACAGAAGAAGAGAGCGCCAGACGCGGCCATCTCTCAAGCCG 2087
QY 1025 alAlaAspLysValLysGluGluSerGlnValProAlaThrGlnThrValGlnArgThrG 1045
DB 2088 TTGCACACAAGGTGAAAGAGAGGTCCAGGTGCTCTCAACCCAGACGTGTGCAGAGAACCG 2147
QY 1045 lySerLysAlaLeuGluLysValGluGluValGluLysSerGluValLeuAlaSerG 1065
DB 2148 GGTCAAAAGCCTGGAGAAAGGTTGAGAGGTAGAGAGGACTCCGAAGTGTCTGGCTTCGG 2207
QY 1065 lylsGluLysAspValMetProLysGlyProValGlnGluAlaGlyAlaGluHisLeuA 1085
DB 2208 AGAAAGAGAAAGACGTTATGCCGAAGAGACCGGTGCAGGAGCTGGAGCTGAGCATCTTG 2267
QY 1085 laGlnGlySerGluThrGlyGlnAlaThrProGluSerLeuGluValProGluValThrA 1105
DB 2268 CACAGGGCTCTGAGACTGGACAGGCTACTCCAGAGAGCCTTGAAGTTCCTGAAGTCACAG 2327
QY 1105 laAspValAspHisValAlaThrCysGlnValIleLysLeuGlnLeuMetGluGlnA 1125
DB 2328 CAGATGACACCATGTCGCCACGTGCCAGTTATCAAGCTCCAGCAGCTGATGGAAACAG 2387
QY 1125 laValAlaProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrProL 1145
DB 2388 CCGTGGCCCTTGAGTCTATCCGAACCTTCACACAGCTGAGACAAATGGNAGCATCTCCCT 2447
QY 1145 euAlaAspSerAspThrAlaAspGlyThrGlnGlnAspGluThrIleAspSerGlnAspS 1165
DB 2448 TAGCAGATTTCAGACACTGCAGATGGGACACAGCAAGATGAACCAATTGACAGCAGGACA 2507
QY 1165 erLysAlaThrAlaAlaValArgGlnSerGlnValThrGluGluGluAlaAlaThrAlaG 1185
DB 2508 GTAAAGCCCATCGAGCTGTGAGCAGCTCACAGGTCAAGAGAAAGAGGCGGTACTGCTC 2567
QY 1185 lnLysGluGluProSerThrLeuProAsnValProAlaGlnGluGluHisGlyGluG 1205
DB 2568 AGAAAGAGAGGCTTCGACACTACCTAATATGTTCCAGCCAGGAGAAACATGGGGAAG 2627
QY 1205 luProGlyArgAspValLeuGluProThrGlnGlnLeuThrAlaAlaValProV 1225
DB 2628 AACCGAAGAGAGATGTTCTTGAACCTACACAGCAAGAGCTTGCTGTGACGCGCTGCCCG 2687
QY 1225 alLeuAlaLysThrGluValGluGlnGluGlyValAspThrProLeuAspGlyGluLysV 1245
DB 2688 TCTGGCAAAAGACTGAGGTGGGTCAAGAGGGTGAGGTTGACTGGTTGGATGGAGAAAAAG 2747
QY 1245 alLysGluGluGlnGluValPheValHisSerGlyProAsnSerGlnLysAlaAlaAspV 1265
DB 2748 TCAAGAAAGAACAGAGAGGTGTTTGTACACTCTGGACCCCAACAGTCAAAAGGCTGCTGATG 2807
QY 1265 alThrTyrAspSerGluValMetGlyValAlaGlyCysGlnGluLysGluSerThrGluV 1285
DB 2808 TGACATATGACAGTGAAGTGATGGGAGTGGCGGGTGTGAGGAAAAGAGAGAGTACTGAAG 2867
QY 1285 alGlnSerLeuSerLeuGluGluGlyGluMetGluThrAspValGluLysGluLysArgG 1305
DB 2868 TGCAGAGTCTTAGCCTGGAGGAGAGAGATGGAATCTGACGTTGAAAAGGAGAAAAAGGG 2927
QY 1305 luThrLysProGluGlnValSerGluGluGlyGluGlnGluThrAlaAlaProGluHisG 1325
DB 2928 AGCAAAAGCCAGAGCAAGTGAGTGAAGAGGTTGAGCAGGAAACAGCCGCTCTCTGAGCATG 2987
QY 1325 luGlyThrTyrGlyLysProValLeuThrLeuAspMetProSerSerGluAlaTrgGlyLysA 1345
DB 2988 AANGGAACCTACGGGAAGCCAGTCTCTGACACTTGACATGCCCCAGCTCAGAGAGGGGAGG 3047
QY 1345 laLeuGlySerLeuGlyGlySerProSerLeuProAspGlnAspLysAlaGlyCysIleG 1365


```

Db 3048 CACTGGGAAGCCTTGGAGGAGCCCTTCTCCAGACCAAGCAAGAGGTTGCTAG 3107
Qy 1365 JuvAIGlnValIGlnSerLeuAspThrValThrGlnThrAlaGluValaValGluValSV 1385
Db 3108 AGGTTCAAGTTCAGAACGCCCTGGACACAAACAGTCACTCAACACAGAGAGCTGGGAAAG 3167
Qy 1385 aLIIGLThrValValIIeserGluThrGluGluSerProGluCysValaIaIaHsEL 1405
Db 3168 TCATAGAAACGGGTGTGATTTCAGAGACAGGTGAAAGTTCAGAGTGTGTAGGTCACACT 3227
Qy 1405 euleProAlaGluLysSerSerAlaThrGlyYHsIstPrThrLeuGlnHsIaIaGlu 1425
Db 3228 TATTACAGCTAGAGAGTCTCTGCAAGCGGTGGCCACTTTCAGCATGAGAGAG 3287
Qy 1425 sPrThrValProLeuGluYProGluSerGlnAlaGluSerIleProIleIleValThrPro 1445
Db 3288 ACACCGTACCCCTGGGGCTGAGTCTCAGGCAAGATCCCAATCCCAATCAATGTAACCTCTG 3347
Qy 1445 lAProGluSerThrLeuHsIspProAspLeuGlnGlyGluIleSerAlaSerGlnArgGlu 1465
Db 3348 CTCCTGAAGGCAACCTTACCTGACCTACAGAGAAATTAAGCGCATCCCAAGAGAGC 3407
Qy 1465 rGserGluGluGluAspLysProAspAlaGlyProAspAlaAspGlyLysGluSerThra 1485
Db 3408 GATCAGAGAGAGAGAGCAAGCCAGATGCTGCTGATGCTGACGGCAAGAGAGATACAG 3467
Qy 1485 lAlIeGluLysValLeuLysAlaGluProGluIleLeuGluLeuGluSerLysSerAsnL 1505
Db 3468 CAATCGAACAAAGTCTCAAGGCTGAACCTGAGACTGACCTGAGACTGAAGCAACA 3527
Qy 1505 YsIIeValLeuAsnValIIeGlnThrAlaValAspGlnHsIaIaArgThrGluThrAlaP 1525
Db 3528 AAGTGTGTCTGAACGATTCAGACAGCCGTTGACCAAGTTCGACCTACAGAAACAGCCC 3587
Qy 1525 rGluThrIstAlaIYrAspSerGlnThrGlnValProAlaCysArgLeuAspSerArg 1545
Db 3588 CCGAAATCATATGCTTATGATTCACAGACCCAGGTTCTGCAATGGCTTGGACAGAGG 3647
Qy 1545 lUpProAsnArgCysTrpThrLysMetLysAspAlaLysMetLysHsIspProValProGlnP 1565
Db 3648 ACCCCCAACAGATGCTGGAACAAATGAATGATGCCAAGATGAACACCCAGTCCCAAC 3707
Qy 1565 rAArgGluAspLeuGlnValLeuThrValLeuGlnAlaIstPrAlaGlnProArgLysCysL 1585
Db 3708 CCAGAGAGAGACTTGCAAGTCTGACCGTCTGAGGCAATGGC--TCAGCTCGGAAATGCT 3765
Qy 1585 euProArgLeuGlnLeuLysAlaProValSerLys 1596
Db 3766 TGCCGGCTTGCAAGTTCGAAGCCCGGTGTCTCAAG 3800

RESULT 11
US-09-902-432-1
; Sequence 1, Application US/09902432
; GENERAL INFORMATION:
; APPLICANT: Irwin H. Gelman
; APPLICANT: Susan G. Jaken
; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
; FILE REFERENCE: A3058-A-FWC-A 070156.0597
; CURRENT APPLICATION NUMBER: US/09/902,432
; CURRENT FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 08/978,277
; PRIOR FILING DATE: 1997-11-25
; PRIOR APPLICATION NUMBER: 08/665,401
; PRIOR FILING DATE: 1996-06-18
; PRIOR APPLICATION NUMBER: 08/635,121
; PRIOR FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 5134
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-09-902-432-1

```

```

Alignment Scores:
Pred. No.: 6,18e-300 Length: 5134
Score: 6005.00 Matches: 1232
Percent Similarity: 97.33% Conservative: 7
Best Local Similarity: 96.78% Mismatches: 27
Query Match: 74.38% Indels: 14
DB: 34 Gaps: 4

US-09-902-432-4 (1-1596) x US-09-902-432-1 (1-5134)
Qy 329 GlnYsThrGluProAlaSerGluGluGlnIupProAlaGluAspThrAspGlnAlaArg 348
Db 2 GAAAGACAGAGCCAGCTCCGAGGAGCAGAGCCGCGAGAAAGACACAGACAGCCAGC 61
Qy 349 LeuSerAlaAspYrGlnLysValaGluLeuProLeuGluAspGlnValaGlyAspLeuGlu 368
Db 62 TTGTACGACAGACTACGAAGAAGGTGACCTGCTTGGAAAGACCAAGTTGTGACCTGGAG 121
Qy 369 AlAserSerGluGluLysCysAlaProLeuAlaThrGluValaPheAspGluLysMetCln 388
Db 122 GCATGTCAGAGAGAGAGAGTGTGCTCTTGGCAACGGAAGTGTGTGATGAGAAAGTGA 181
Qy 389 AlAsIsgInGluValaIaIaGluValaHsIstPrThrValaGluLysThrGluGluGlu 408
Db 182 GCCCACCAAGAAATGTTGTCAGAGAGTCCACGTGACACCTGGAGAAACAGAGAGAG 241
Qy 409 GlnGlyGlyGlyGlyGluAlaGluGlyGlyValaValaGluGlyThrGlyGluSerLeu 428
Db 242 CAGGAGAGAGAGAGAGAGCTGAAGAGGGGCTGTGTGTAAGAAACAGAGAAATCCTTG 301
Qy 429 ProProGluYsIleValaGluProGlnGluValaProGlnGluAlaGluProAlaGluGlu 448
Db 302 CCCCTGAAGAACTGGCTGAGCCCCAGAGAGTCCCCAGGAAGCTGAGACCTGTAGAGAG 361
Qy 449 LeuMetLysSerArgGluMetCysValSerGlyGlyAspHsIstPrGlnLeuThrAspLeu 468
Db 362 CTGATGAAGACAGAGAGATGTGTCTCTGAGAGAGACCACTCAACTGACAGACCTTA 421
Qy 469 SerProGluGluLysThrLeuProLysHsIspProGluGlyIleValSerGluValaGluMet 488
Db 422 AGTCTGAAGAGAGAGACCTCTGCCCAACACCAAGAGCATGTCAGAGAGTGAAGATG 481
Qy 489 LeuSerSerGlnGlnLysIleLysValaGlnLysSerProLeuLysLysLeuPheSerSer 508
Db 482 CTGTCTCTCAGGAAGAAATCAAGGTACAGGAACTCCCTTGAAGAACTCTTCACTGTGC 541
Qy 509 SerGlyLeuLysLysLeuSerGlyLysLysGlnLysGlyLysArgGlyGlyGlyYasp 528
Db 542 TCAGGCTTAAAGAACTGTCTGGAGAGAAAGAGAAAGGAGAAACAGAGAGTGGGGAGAGC 601
Qy 529 GlnIupProGlyGluYrGlnHsIstPrThrGluSerProGluSerAlaAspGluGln 548
Db 602 GAAAGACCTGAGGAATACCAACACATTCACACCGAATCCCAAGAGAGTGTGATGAGAG 661
Qy 549 LysGlyGluSerSerAlaSerSerProGluGluProGluGluThrCysLeuGluLys 568
Db 662 AAGGAGAGAGTCTGCGTCTGCTCCCGCAGAGAGCTTGAGAGACCACTGTCTGAGAA 721
Qy 569 GlyProLeuGluAlaProGluAspGlyGluAlaGluGluGluGlyThrThrSerAspGlyGlu 588
Db 722 GGGCCGCTGAAAGACCCAG--GATGGGAAGGTGAGAAAGAACTACTTCTG--GGAAG 778
Qy 589 LysLysArgGluGlyIleThrProTrpAlaSerPheLysLysMetValThrProLysLys 608
Db 779 AAGAGAGAGAG--ATCATCTCCCTGGGCATCTTCAAAAAGATGTGACACCCAAAGAA 836
Qy 609 ArgValaArgArgProSerGluSerAspLysGluGluGluLysValaLysSerAla 628
Db 837 CGGTC--CGAAGACTTTCGAGAGTGAAGAGAGAGAGAGCTGGAAGAGTCAAGAGCC 895
Qy 629 ThrLeuSerSerThrAspSerThrValaSerGluMetGlnAspGluValaLysThrValaGly 648

```

Db 896 ACCTTGCTCCACTGATACAGTCAGTGTGAGAAATGCAAGATGAAGTCAAACTGTTGGT 955
Qy 649 GluGluGlnLysProGluGluProLysArgValAspThrSerValSerTrpGluAla 668
Db 956 GAGGAACAAAAGCCAGAGGAACCAAGCGTAGGTGATACCTTCAGTGTCTTGGGAAGCA 1015
Qy 669 LeuLeuCysValGlySerSerLysLysArgAlaArgLysAlaSerSerSerAspAspGlu 698
Db 1016 CTGATTGTGTGCGGATCATCCAGAGAGAGCAAGGAAGGCATCTCTTCAGATATAAGA 1075
Qy 689 GlyGlyProArgThrLeuGly-GlyAspSerHisArgAlaGluGluAlaSerLysAspLys 708
Db 1076 ---GGGCCCAAGACATCGGAGGGGACAGTCAAGAGCAGAGAGGCCACGAAAGACAA 1132
Qy 708 sGluAlaGlyThrAspAlaValProAlaSerThrGlnGlnAspGlnAlaGlnGlySe 728
Db 1133 AGAAGCCGA-ACAGACGCTGTTCTGCCAGCACCCAGGAGCAGGACCAAGCGCAAGGAAG 1191
Qy 728 rSerSerProGluProAlaGlySerProSerGluGlyGluGlyValSerThrTrpGluSe 748
Db 1192 TTCTCTCACCCGAGCCAGCGGAAAGCCCTTCCGAAGGGGAGGTGTCTCCACTTGGGAGTC 1251
Qy 748 rPheLysArgLeuValThrProArgLysLysSerLysSerLysLeuGluGluLysAla-- 767
Db 1252 ATTTAAAGATTAGTCACTCCAAAGAAAATAATCAAGTCAAACTCGAAGAGAAAGA-AG 1310
Qy 768 ----GluAspSerSerVal----GluGlnLeuSerThrGluLeuGluProSerArgGluL 785
Db 1311 CCGGAAGGACTCTAGTTGTAGGAGCAGGTGTCTCACTGAGATCGAACCCGTGTAGAGAGA 1370
Qy 785 uSerTrpValSerIleLysLysPheLysProGlyArgArgLysArgAlaAspGlyLys 805
Db 1371 ATCTTGGGTTTCCATTAAAGAAATTCATCCCGGACGCGGAAAGAAAGGCGAGATGGGAA 1430
Qy 805 s-GlnGluGlnAlaThrValGluAspSerGlyProValGluLleAsnGluAspAspProA 825
Db 1431 GGCAAGAACAAAGCCACTGTGGAAAGACTCAGGGCCAGTGGAGATAAATGAGAGCAGCCTG 1490
Qy 825 snValProAlaValAlProLeuSerGluTyrAsnAlaValGluArgGluLysMetGluA 845
Db 1491 ATGTCCAGCCGTCGTCTCTCTCTGAGTATGATGATGATGAGAGGGGAGAGATGGAG 1550
Qy 845 laGlnGlyAsnThrGluLeuProGlnLeuLeuGlyAlaValTyrValSerGluGluLeuS 865
Db 1551 CCAGGGGAAATGCGGAGCTGCCAGCTGTCTGGGGCTGTGTA---GTGTCGAGGAGCTCA 1607
Qy 865 erLysThrLeuValHisThrValSerValAlaValIleAspGlyThrArgAlaValThrS 885
Db 1608 GTAAGACTCTGGTCCACACTGTGAGTGTGCGAGTCAATTGATGGGACAGGGCAGTCACCA 1667
Qy 885 erValGluArgSerProSerTrpIleSerAlaSerValThrGluProLeuGluHist 905
Db 1668 GTGTGGAAGCGGTCTCTCTCTGGGATATCCGTTCCGTAACAGAAACCTCTTGAACACA 1727
Qy 905 hrAlaGlyGluAlaMetProProValGluGluValThrGluLysAspIleIleAlaGluG 925
Db 1728 CACGGGAGAGGCATGCCACCTGTTGAAGAGGTCACTGAAAAAGACATCATTTGCAGAAAG 1787
Qy 925 luThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHisAspAspMetValT 945
Db 1788 AAACTCTGTGCTCACCCAGACGTTACCAGAGGGTAAAGATGCCCATGACGACATGGTCA 1847
Qy 945 hrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeuA 965
Db 1848 CCAGTGNAGTGGATTTCACCTCAGAGAGTGTGACAGCCACAGAGACCTCNAGAGCTCTCC 1907
Qy 965 rgThrGluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSerAlav 985
Db 1908 GTACTGAAGAAGTTACCGAAGCATCGGGGCCGGAAGAGACCAAGACATGCTGTCTCCGAG 1967
Qy 985 aIsrGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGluValG 1005
Db 1968 TTTTCCAGCTGACTGATCTCCCAAGACCAACAGAGGAAGCCACCCAGTTTCAGAGGTAG 2027

Qy 1005 luSerGlyValLeuAspThrGluGluGluArgGlnThrGlnAlaIleLeuGlnAlav 1025
Db 2028 AGGGTGTGTGTAGATACAGAAAGAGGAGCGCCAGACGCGGCACTCCTCCAGCGCG 2087
Qy 1025 alAlaAspLysValLysGluGluSerGlnValProAlaThrGlnThrValGlnArgThrG 1045
Db 2088 TTGCAGACAAAGGTGAAAGAGGAGTCCCAGTGCCTGCAACCCAGACTGTGCGAGAAACGG 2147
Qy 1045 lySerLysAlaLeuGluLysValGluGluValGluGluAspSerGluValLeuAlaSerG 1065
Db 2148 GGTCAAAAGCATCTGGAGAGGTTGAGGAGGTAGAGGAGACTCCGAGAGTGTGGCTTCGG 2207
Qy 1065 luLysGluLysAspValMetProLysGlyProValGlnGluAlaGlyAlaGluHisLeuA 1085
Db 2208 AGAAAGAGAAGACGCTTATGCCGAAAGGACCCTGTCAGGAAGCTGGAGCTGAGCATCTTG 2267
Qy 1085 laGlnGlySerGluThrGlyGlnAlaThrProGluSerLeuGluValProGluValThrA 1105
Db 2268 CACAGGGCTCTGAGACTGGACAGGCTACTCCAGAGAGCCTTGAAGTTCCTGAAGTCACAG 2327
Qy 1105 laAspValAspHisValAlaThrCysGlnValIleLysLeuGlnGlnLeuMetGluGlnA 1125
Db 2328 CAGATGTAGACCATGTGCGCACGTGCCAGGTTATCAAGCTCCAGCAGCTGATGGAAACAGG 2387
Qy 1125 laValAlaProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrProL 1145
Db 2388 CCGTGGCCCTCGAGTCATCCGAAACCTTCACAGACAGTGCAGCAAAATGGAAGCAGCTCCCT 2447
Qy 1145 euAlaAspSerAspThrAlaAspGlyThrGlnGlnAspGluThrIleAspSerGlnAsps 1165
Db 2448 TAGCAGATTTCAGACACTGCAGATGGGACACAGCAAGATGAAACCATTTGACAGCAGACA 2507
Qy 1165 erLysAlaThrAlaAlaValArgGlnSerGlnValThrGluGluAlaAlaThrAlaG 1185
Db 2508 GTAAGCCACTGCAGCTGTGAGCGAGTCACAGTTCACAGAAAGAGGGCGGTACTGTCTC 2567
Qy 1185 lnLysGluGluProSerThrLeuProAsnAsnValProAlaGlnGluGluHisGlyGluG 1205
Db 2568 AGAAAGAGGAGCCTTCGACACTACCTAATATGTTCCAGCCAGGAGAAACATGGGGAAG 2627
Qy 1205 luProGlyArgAspValLeuGluProThrGlnGlnGluLeuThrAlaAlaAlaValProV 1225
Db 2628 AACCCAGGAAGAGATGTTCTTGAACCTTACACAGCAAGAGCTTGTGTCGACCGCTCCCG 2687
Qy 1225 alLeuAlaLysThrGluValGlyGlnGluGlyValAspTrpLeuAspGlyGluLysV 1245
Db 2688 TCTGGCAAAAGACTGAGGTGGGTCAAGAGGTGAGGTTGACTGTGGATTGGAGAAAAAG 2747
Qy 1245 alLysGluGluGlnGluValPheValHisSerGlyProAsnSerGlnLysAlaAlaAspV 1265
Db 2748 TCAAGAAGAAACAGGAGGTGTTTGTACACTCTGGACCCCAACAGTCAAAAGGCTGCTGATG 2807
Qy 1265 alThrTyrAspSerGluValMetGlyValAlaGlyCysGlnGluLysGluSerThrGluV 1285
Db 2808 TGACATATACAGTGAAGTGAATGGGAGTGGCGGGTGTCCAGGAAAGGAGAGTACTGAAG 2867
Qy 1285 alGlnSerLeuSerLeuGluGlyGluMetGluThrAspValGluLysGluLysArgG 1305
Db 2868 TCAGAGTCTTAGCTTGGAGGGGAGAGATGGAACACTGACGTTGAAAAAGGAGAAAAAGG 2927
Qy 1305 luThrLysProGluGlnValSerGluGlyGluGlnGluThrAlaAlaProGluHisG 1325
Db 2928 AGCAAAAGCCAGACNAGTGAAGAGGTGAGCAGGAAACAGCCGCTCTCTGACATG 2987
Qy 1325 luGlyThrTyrGlyLysProValLeuThrLeuAspMetProSerSerGluArgGlyLysA 1345
Db 2988 AAAGGAACCTACGGGAAGCCAGTCTTGCACACTTGACATGCCAGTCCAGAGAGGGGAAGG 3047
Qy 1345 laLeuGlySerLeuGlyGlySerProSerLeuProAspGlnAspLysAlaGlyCysIleG 1365
Db 3048 CACTGGGAAGCCTTGAGGAGAACCTTCTCTCCAGACCAAGCAAGCAGGTTGCATAG 3107

QY 323 ValAspGluGluGluLysGluLysThrGluProAlaSerGluGlu----- 337
Db 1185 GTAGACAGAGAAGAACGGAAGGCGAGGTTCCCTCCGAGAAATGACCGCTCCGAG 1244
QY 338 -----GlnGluProAlaGluAspThrAspGlnAlaArgLeuSerAlaAspTyr 353
Db 1245 CAAGCCACCCACAGGAGCGGAGAGAGTCCCAAGAGTCCCAAGCGCTTATCAGCTGATAT 1304
QY 354 GluLysValGluLeuProLeuGluAspGlnValGlyAspLeuGluAlaSerSerGluGlu 373
Db 1305 GAGAAAGTTGAGCTGCCCTCAGAGGAGCAAGTCACTAGTGGCTCGCAGGAGACCTTCTGAAGAG 1364
QY 374 LysCysAlaProLeuAlaThrGluValPheAspGluLysMetGluAlaHisGln--Glu 392
Db 1365 AAACCTGCTCCGTGGCGAGAGAGTGTTCATGAGAAATAGAAAGTCCACCAAGAGAG 1424
QY 393 ValValAlaGluValHisValSerThrValGluLysThrGluGluGlnGlyGly 412
Db 1425 GTTGTGGCCGAGTCCACGTCAGCACCGTGGAGGAGAGAACCCGAGAGAGCAG----- 1475
QY 413 GlyGluAlaGluGlyGlyValValGluGlyThrGlyGluSerLeuProGluLys 432
Db 1476 -----AAAAAGGAGTGGAAAGAACAGCAGCGGTCTGTGCCAGCTGAAGAA 1520
QY 433 LeuAlaGluProGlnGluValProGlnGluAlaGluProAlaGluLeuMetLysSer 452
Db 1521 TTGGTTGGAAATGGATGAGAACCTTCAGGAAGCCGAACTGCCAAGGAGCTGGTGAAGCTC 1580
QY 453 ArgGluMetCysValSerGlyGlyAspHisThrGlnLeuThrAspLeuSerProGluGlu 472
Db 1581 AAAGAAACGTGTGTTCCGAGAGGAGCCCTACAGGGAGCTGACCTCAGTCTGTATGAG 1640
QY 473 LysThrLeuProLysHisProGluGlyLeValSerGluValGluMetLeuSerSerGln 492
Db 1641 AAGTGTGTGCCAAACCCCGGAAGGCGTTGTGAGTGGAGTGGAAATGCTGTCAACAG 1700
QY 493 GluArgLysValGlnGlySerProLeuLysLysLeuPheSerSerSerGlyLeuLys 512
Db 1701 GAGAGAAATGAAGTGCAGGGAAGTCCATAAGAAAGCTTTTACCAAGCTGGCTTAAAA 1760
QY 513 LysLeuSerGlyLysLysGlnLysGlyLysArgGlyGlyGlyGlyAspGluGluProGly 532
Db 1761 AGCTTTCTGGAAAGAAACAGAAAGGGAAGA--GGAGAGGAGACGAGGAATCAGGG 1817
QY 533 GluTyrGlnHisLeHisThrGluSerProGluSerAlaAspGluGlnLysGlyGluSer 552
Db 1818 GAGCACACTCAGGTTCCAGCGGATCTCCGACAGCAGGAGGAGCAAAAGGGCGAGAGC 1877
QY 553 SerAlaSerProGluGluProGluGluThrThrCysLeuGluLysGlyProLeuGlu 572
Db 1878 TCTGCCCTCATCCCTGAGGAGCCGAGGAGATCACGTTGTCTGAAAAAGGGCTTAGCCGAG 1937
QY 573 AlaProGlnAspGlyGluAlaGluGluGlyThrThrSerAspGlyGluLysLeuArgGlu 592
Db 1938 GTGCAGCAGGATGGGAAGCTGAAGAAGAGCTACTTCCGATGGAGAGAAAGAAAGAGAA 1997
QY 593 GlyLeuThrProTyrAlaSerPheLysMetValThrProLysLysArgValArgArg 612
Db 1998 GGTGTCACTCCCTGGGCATCATTCAAAAGATGTTGAGCGCCCAAGAGCGTTAGACGG 2057
QY 613 ProSerGluSerAspLysGluGluGluLeuLysValLysSerAlaThrLeuSerSer 632
Db 2058 CTTTCGAAAAGTGATAAAGAGAGATGAGCTGCAAGGTCAAGAGCGCTACCTGTCTTCC 2117
QY 633 ThrAspSerThrValSerGluMetGlnAspGluValLysThrValGlyGluGluLys 652
Db 2118 ACCGAGACACAGCCTCTGAAATCCAGAAAGAAATGAAGGGAGCGGTGGAGAGCCAAAG 2177
QY 653 ProGluGluProLysArgValAspThrSerValSerTyrGluAlaLeuLeuLysVal 672
Db 2178 CCGAAGAACCAAGCCAGAGTGGATACCTCAGTATCTTGGGAGCTTTAATTTGTGTG 2237
QY 673 GlySerSerLysLysArgAlaArgLysAlaSerSerSerAspGluGlyGlyProArg 692

Db 2238 GGATCATCCAAAGAAAAGACAGAGAGGGTCTCTTCTGATGAGGAAGGGGACCAAAA 2297
QY 693 ThrLeuGlyGlyAspSerHisArgAlaGluAlaSerLysAspLysGluAlaGlyThr 712
Db 2298 GCAATGGAGGAGACCACCAGAAAGCTGATGAGCCGGAAGAACCAAGAGAGCGGAC 2357
QY 713 AspAlaValProAlaSerThrGlnGluGlnAspGlnAlaGlnGlySerSerProGlu 732
Db 2358 GACGGATCTTCTGTTCCCAAGAACATGATCCAGGGCAGGGAAGTCTCTCCCGGAG 2417
QY 733 ProAlaGlySerProSerGluGlyGlyValSerThrTrpGluSerPheLysArgLeu 752
Db 2418 CAAGCTGGAGAGCCCTACCGAAGGGAGGGGTTTCCACCTGGGAGTCATTTAAAGGTTA 2477
QY 753 ValThrProArgLysLysSerLysSerLysLeuGluGluLysAlaGluAsp----- 769
Db 2478 GTCAAGCCCAAGAAAATCAAGTCCAGCTGGAGAGAAAGCGAAGACTCCATAGCT 2537
QY 770 ---SerSerValGluGlnLeuSerThrGluLeuGluProSerArgGluGluSerTyrVal 788
Db 2538 GGTCTGGTGTAGAACATTCCTCCTCAGACACTGAACCCCGTAAAGAAATCTCTGGGTC 2597
QY 789 SerLeuLysPheLeuProGlyArgArgLysLysArgAlaAspGlyLysGlnGluGln 808
Db 2598 TCAATCAAGAGGTTTATTCCTGACCAAGAGAAAGGCCAGATGGGAAACAAAGACAA 2657
QY 809 AlaThrValGluAspSerGlyProValGluLeuAsnGluAspAspProAsnValProAla 828
Db 2658 GCCCTGTTGAAGACGAGCGGCGCCACAGGGGCCAACGAAGATCCTCTGTATGTCGCGCC 2717
QY 829 ValValProLeuSerGluTyrAsnAlaValGluArgGluLysMetGlu-----Ala 845
Db 2718 GTGGTCCCTCTGTCTGATGATGCTGTAGAAAGGAGAGAAATGGAGGCACAGCAAGCC 2777
QY 846 GlnGlyAsnThrGluLeuProGlnLeuLeuGlyAlaValTyrValSerGluGluLeuSer 865
Db 2778 CAAAAAGCGCAGACGCGCCGAGAGAGGCGACCACTGAGGTGTCCAAGAGCTCAGC 2837
QY 866 LysThrLeuValHisThrValSerValAlaValIleAspGlyThrArgAlaValThrSer 885
Db 2838 GAGAGTCAGGTTCTATGATGGCAGCAGCTGTCTGACGGACGAGGCGGCTACCAT 2897
QY 886 ValGluGluArgSerProSerTyrIleSerAlaSerValThrGluProLeuGluHisThr 905
Db 2898 ATTGAAGAAAGGTCTCTCTTGGATATCTGCTTCAAGTACAGAACCTCTTGAACAAGTA 2957
QY 906 AlaGlyGluAlaMetProProValGluGluValThrGluLysAspIleAla---Glu 924
Db 2958 GAAAGCTGAAGCCGCACTGTTAACTGAGGAGGTATTGGAAAGAGAGTAATTTGAGAAGAA 3017
QY 925 GluThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHisAspAspMetVal 944
Db 3018 GAACCCCGCAGGTTACTGAACCTCTGCCAGAGAACAGAGAGCGCCGCGGCGACCGTC 3077
QY 945 ThrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeu 964
Db 3078 GTTAGTGAGCGGAATTGACCCCGGAGCTGTGACAGCTGCAGAAACTGCAGGCGCAT 3137
QY 965 ArgThrGluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSerAla 984
Db 3138 GGTGCGGAAGAGAAACCGAAGCATCTGTCTGCTGAAGAGACCAAGAAATGGTGTGAGCA 3197
QY 985 ValSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGluVal 1004
Db 3198 GTCTCCAGTTAAACCCACTCCCGACACACACAGAGAGGCGCACTCCGGTGCAGGAGGTG 3257
QY 1005 GluSerGlyValLeuAspThrGluGluGluGluGluArgGlnThrGlnAlaLeuLeuGlnAla 1024
Db 3258 GAAGGTGGGTACTCTCAGATAGAGAGAGAGAGGCGGACTCAAGAGGTCTCTCCAGGCA 3317
QY 1025 ValAlaAspLysValLysGluGluSerGlnValProAlaThr----- 1038

```

Db 3318 GTGGCAGAAAAAGTGAAGAGGAAATCCAGCTGCTGGACCGCGTGGCCAGAAAGATGTC 3377
QY 1039 ---GlnThrValGlnArgThrGlySerIysAlaLeuGluLysValGluGluValGluGlu 1057
Db 3378 CTTACGCTGCTGCACAGAGCA-----GAGGCAGAAAGACCCAGAAAG 3419
QY 1058 AspSerGluValLeuAlaSerGluLysGlyLysAspValMetProLysGlyProValGln 1077
Db 3420 CAGGCTGAAGCCCTCGGCTGTGAAGAAAGAGACGATGTATGTGAAGTATAGTCTCAG 3479
QY 1078 GlnAlaGluValAlaGluHisLeuAlaGlnIleGlnIleGlnIleGlnIleGlnIleGlnIle 1097
Db 3480 GAGGCAAAAACCTGAGCCTTTTACACAGGGAAGGTGTGGGCGACACCCCGAGAAAC 3539
QY 1098 LeuGlu---ValProGluValThrAlaAspValAspHis-----ValAlaThrCys 1113
Db 3540 TTTGAAAAAGCTCTCAAGTACACAGAGACATAGATTCAGTGAAGCTTTAAACCTTGT 3599
QY 1114 GlnVal-----IleLysLeuGlnGlnLeu---MetGluGlnAlaValAla 1127
Db 3600 CAAGCCGAAACCTTAGCTGGGTGTAATAATCAGAGATGTGATGGAACAGGCTATCCCC 3659
QY 1128 ProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrProLeuAlaAsp 1147
Db 3660 CCGTACTCGGTGGAACCCCTTACAGACAGTGAAGTGTGATGAAAGCAGCCCTGAGCCGAC 3719
QY 1148 SerAspThrAlaAspGlyThrGlnIleAspGluThrIleAspSerGlnAspSerIysAla 1167
Db 3720 TTTGACGACCCAGGACCAACCCAGAAAGCGAATGTGGAATTCATGAGAGGATGAG 3779
QY 1168 ThrAlaAlaValArgGlnSerGlnValThrGluGluGluAlaAlaThrAlaGlnLysGlu 1187
Db 3780 GTGGATCTGGTATCCAGTACAGGGGCGACAGAGCAAGAGCGATTCCTTCACAGAAAGAG 3839
QY 1188 GluProSerThrLeuProAsnAsnValProAlaGlnGluGlnHisGlyGluGluProGly 1207
Db 3840 AGGCTCCAGCA---CCTTCAGTTTGTGTTCACAGAAAGAACTAAAGAAACATATAAG 3896
QY 1208 ---ArgAspValLeuGluProThrGlnGlnGlnIleThrAlaAlaAlaValProValLeu 1226
Db 3897 ATGGAAGACACTCTAGAGCATACAGATAAAGAGGTGTGAGTGAAGCTGATTCATCTCG 3956
QY 1227 AlaLysThrGluValGlyGlnGlnGluGluValAspTrpLeuAspGlyGluValLys 1246
Db 3957 TCAAAGACTGAGGGGACTCA-----GAGGCTGACCACTATGCTGATGAGAAACCAAA 4010
QY 1247 -----GlnGluGlnGluValPheValHisSerGly-----ProAsnSer 1259
Db 4011 GACGTACCATTTTTCGAAGGACTTGAGGGGTCTATAGACACAGGCAATAACAGTCACTCG 4070
QY 1260 GlnLysAlaAlaAspValThrTrpAspSerGluValMetGlyValAlaGluCysGlnGlu 1279
Db 4071 GAAAAGGTACATGAAGTTCCCTTAAAGTGAAGGACAGAAAGAACTGAATGTAATAAG 4130
QY 1280 LysGluSerThrGluValGlnSer-----LeuSerLeuGluGluGluGlu 1294
Db 4131 GATGATGCTCTTGAAGTCAAGAGTCAAGCTAAGTCTCTCCATCCCGCTGGAGAGAGAG 4190
QY 1295 MetGluThrAspValGluLysGluLysAspGluThrLysProGluGlnValSerGluGlu 1314
Db 4191 ATGGTATGTTCAAGTCAAGAGGAGAAACAGAAAGCAGAGCCAAACCATGATGATGAAGAG 4250
QY 1315 Gly---GluGlnGluThrAlaAlaProGluHisGluGluThrTrpGlyLysProValLeu 1333
Db 4251 AAGCTTGACACAGAAACACACTGTTACCGTATCGAAAGAGTCAAGTGAAGAGCTCCACAG 4310
QY 1334 ThrLeuAspMetProSerSerGluLysGlyLysAlaLeuGlySerLeuGlySer--- 1352
Db 4311 ACAGTGAATGTCCCATCATATAGATGGGCAAGAGTCAAGAGTTCGAAAGAGACCCCT 4370
QY 1353 ProSerLeuProAspGlnAspLysAlaGlyCysIleGluValGlnValHisLeuAsp 1372
Db 4371 CCTCCCTGCTAGGTCAAGAGAGGACGATGACCAAAATTCATAGTACAGAGCTCTGAG 4430

```

```

QY 1373 ThrThrValThrGlnThrAlaGluAlaValGluLysValIle-----GlnThrValVal 1390
Db 4431 GCATCATCTACCTTAACACGCGCTGCACAGAGAGAAAGAGTCTTAAGAGAACTGCCAAC 4490
QY 1391 IleSerGluThrGlnGluSerProGluCysValGlnAlaHisLeuLeuProAlaGluLys 1410
Db 4491 ATTTTGAACACAGGTGAACAGTGTGAGGCTTCAGAGTGCACATTTAGTTCTGGAAGAGAA 4450
QY 1411 SerSerAlaThrGlnGluHisIleThrLeuGlnHisAlaGluAspThrValProLeuGly 1430
Db 4551 TCCTGTGAAAAAATGAAGACTTTTGCCCTCATCCAGGGAAGATGCTGTGCCACAGG 4610
QY 1431 ProGluSerGlnAlaGluSerIleProIleIleValThrProAlaProGluSerThrLeu 1450
Db 4611 CCGCAGTCTCAGGCAAAATTCACACACAGTGAATGATCTGCTACTACCAAGAAAGGCTTA 4670
QY 1451 HisProAspLeuGlnGluIleSerAlaSerGlnArgGluArgSerGluGluLysP 1470
Db 4671 AGTTCGACCTGGAAAGAGAGAAACCAATCATCTGAAGTGAAGTCAAGATGAAGTCAAT 4730
QY 1471 LysProAspAlaGlyProAspAlaAspGlyLysGluSerThrAlaIleGluValLeu 1490
Db 4731 GAGCAGTGTGCTTGCACAGAGGTC-----AAAGTGAAGTGAAGATTTGAGAGAGATTTA 4784
QY 1491 LysAlaGluProGluIleLeuGlnLeuGlnSerLysSerAsnLysIleValLeuAsnVal 1510
Db 4785 GAGCCTGAAAAATGGGATTTTGAACCTTGAGACCAAAACAGATTAACCTGTCACAAAACATC 4844
QY 1511 IleGlnThrAlaValAlaAspGlnPheAlaArgThr---GluThrAlaProGluThrHisAla 1529
Db 4845 ATCCAGACAGCCGTGACACAGTGTGATCAGTACAGAAAGAACCCACCGAAATGTTGACG 4904
QY 1530 TyrAspSerGlnThrGlnValProAlaCysArgLeuAspSerArgGluProAsnArgCys 1549
Db 4905 TGTGAGTTACAGACACAGAGCTCAGGTGATTAAGCTGACACCGAGAGCGTGCAGAGAA 4964
QY 1550 TrpThrLys-----MetLysAspAla-LysMetLysHisProValProGlu 1564
Db 4965 ACGGAGAAAGAGAGAGAGAACTCAGGCTCTGCACAGAGATGAAGAACACCATTTCTTCA 5024
QY 1564 nProArgGlu-----AspLeuGlnValLeuThrValLeuGluAla 1577
Db 5025 GCCAAAGAGAGTCAAGTCAACCGCAGTGGACCAAGCAATTCGATATTTCCAAAGAC 5084
QY 1577 aTrpAlaGlnProArgLysCysLeuProArgLeuGlnLeuLysAlaProVal 1594
Db 5085 ATGATGAAAGCTTCAGAAAGACCATGACTGTGATGATGAAGTTCACATG 5136

RESULT 13
US-60-278-232-10143
; Sequence 10143, Application US/60278232
; GENERAL INFORMATION:
; APPLICANT: Morris, MacDonald
; APPLICANT: Lal, Preeti
; TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using
; TITLE OF INVENTION: Polynucleotide Sequence Databases, and Single Nucleotide
; FILE REFERENCE: GX-0011 P
; CURRENT APPLICATION NUMBER: US/60/278, 232
; NUMBER OF SEQ ID NOS: 2001-03-30
; SOFTWARE: PERL Program
; SEQ ID NO 10143
; LENGTH: 6603
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 445786.12
US-60-278-232-10143

```

Alignment Scores:

Pred. No.: 1,43e-213 Length: 6603
 Score: 4333.50 Matches: 968
 Percent Similarity: 69.07% Conservative: 191
 Best Local Similarity: 57.69% Mismatches: 409
 Query Match: 53.68% Indels: 110
 DB: 71 Gaps: 33

US-09-902-432-4 (1-1596) x US-60-278-232-10143 (1-6603)

QY	1	MetGlyAlaGlySerSerThrGluGlnArgSerProGluGln---ProAlaGlySerAsp	19
DB	193	ATGGGCGCCGGAGCTCCACCGAGCAGCGAGCCGCGGAGCAGCGCCCGGAGGAGCTCC	252
QY	20	ThrProSerGluLeuValLeuSerGlyHisGlyProAlaAlaGluAlaSer---GlyAla	38
DB	253	ACGCCGCTGAGCCGAGCCAGCGCGCGCGCGCCCTCGCGCGCGCGCGCCAGACACC	312
QY	39	AlaGlyAspProAlaAspAla-----AspProAlaThrLysLeuProGlnLysAsnGly	56
DB	313	ACCGCGGACCCCGCCAFCGCTGCTCGGACCCCGCCACCAAGCTCTTACAGAAAGATGGT	372
QY	57	GlnLeuSerSerValAsnGlyValAlaGluGlnGlyAspValHisValGlnGluGluAsn	76
DB	373	CAGCTGTCCACCATCATGCGGTAGCTGACCAAGATGAGCTCAGCTCCAGGAGGGTGAC	432
QY	77	GlnGluGlyGln-----GluGlu	82
DB	433	CTAAATGGCCAGAAAGAGAGCCCTGAACGGTCAAGGAGCCCTAAACAGCCAGGAGGAAGAA	492
QY	83	GluValValAspGluAspValGlyGlnArgGluSerGluAspValArgGluLysAspArg	102
DB	493	GAAGTCAATGTCACAGAGGTGGACAGAGAGACTCTCAAGATGTGAGCAAAAGAGACTCC	552
QY	103	ValGluGluMetAlaAlaAsnSerThrAlaValGluAspIleThrLysAspGlyGlnGlu	122
DB	553	GATAAAGAGATGGCTACTAAGTCAGCGGGTGTTCACGACATCACAGATGATGGCGAGAG	612
QY	123	GluThrSerGluIleLeuGluGlnIleProAlaSerGluAsnValGluGluMetVal	142
DB	613	GAGACACCCGAATAATCGAAGATTCCTTCTTCAGAAAGCAATTTAGAAAGAGCTAACCA	672
QY	143	GlnProAlaGluSerGlnAlaAsnAspValGlyPheLysLysValPheLysPheValGly	162
DB	673	CAACCCACTGAGTCCCAAGGCTAATGATATGGAATTTAAGAAGGTGTTTAAAGTGTGTTGGC	732
QY	163	PheLysPheThrValLysLysAspLysAsnGluLysSerAspThrValGlnLeuLeuThr	182
DB	733	TTTAAATTCATCTGTGAAAAAGGATAACAGAGAGAGCCTGACACTGTCCAGCTACTCACT	792
QY	183	ValLysLysAspGluGlyGluGlyValAlaGluAlaSerValGlyAlaGlyAspHisGlnGlu	202
DB	793	GTGAAGAAAGATGAGGGGAGGAGCA-----CGAGGGCTGGCGACCCAGGAC	843
QY	203	ProSerValGluThrAlaValGlyGluSerAlaSerLysGluSerGluLeuLysGlnSer	222
DB	844	CCAGAGCCTT-----GGGGCTGGAGAGCAGCATCCAAAGAAAGCGAACCAACATCT	897
QY	223	ThrGluLysGlnGluGlyThrLeuLysGlnGlnSerSerThrGluIleProLeuGln	242
DB	898	ACAGAGAAACCCGAAGAGACCCCTGAACGGTGAAGCAAGCCAGCAGCAAAATTTCTCCCCCA	957
QY	243	AlaGluSerAspGlnAlaAlaGluGluAlaLysAspGluGlyGluLysGlnGlu	262
DB	958	GCCGAATCTGGCCNAGCAGTG---GAGGAATGCNAGAGGAGGAGGAAGAAACAAGAA	1014
QY	263	LysGluProThrLysSerProGluSerProSerSerProValAsnSerGluThrThrSer	282
DB	1015	AAAGAACCTAGCAAGTCTGCAGAAATCTCCGACTAGTCCCGTGCACAGTGAACAGGATCA	1074
QY	283	SerPheLysLysPhePheThrHisGlyTrpAlaGlyTrpArgLysLysThrSerPheLys	302
DB	1075	ACCTTCAAAAAATTTCTCACTCAAGGTTGGCGCGCTGGCGGCAAAAAAGACCAAGTTTCAGG	1134

QY	303	LysSerLysGluAspAspLeuGluThrAlaGluLysArgLysGluGlnGluAlaGluLys	322
DB	1135	AAGCCGAAGGAGGATGAAGTGGAAAGCTTCAGAGAAAGAAAGAAAGAAAGAAAGAA	1194
QY	323	ValAspGluGluGluLysGluLysThrGluProAlaSerGluGlu-----	337
DB	1195	GTAGACACAGAAAGAACCGAAAGGAGAGGTTGCTCCGAGAAACTGACCGCTCCGAG	1254
QY	338	-----GlnGluProAlaGluAspThrAspGlnAlaArgLeuSerAlaAspTyr	353
DB	1255	CAAGCCACCCACAGGAGCGGAGAAAGTGCACAGCCCGGTTATCAGCTGAATAT	1314
QY	354	GluLysValGluLeuProLeuGluAspGlnValGlyAspLeuGluAlaSerSerGluGlu	373
DB	1315	GAGAAAGTTGAGCTGCGCTCAGAGGAGCAAGTCAGTGGCTCGCAGGAGACCTTCTGAAGAG	1374
QY	374	LysCysAlaProLeuAlaThrGluValPheAspGluLysMetGluAlaHisGln---	392
DB	1375	AAACCTGCTCCGTTGGCGACAGAAAGTGTGTGATGAGAAATAGAAATCCACCAAGAGAG	1434
QY	393	ValValAlaGluValHisValSerThrValGluLysThrGluGluGlnGlyGlyGly	412
DB	1435	GTGTGCGCGAGTCCACGTCAGCACCGTGGAGGAGAGAACCGAAAGAGAG---	1485
QY	413	GlyGluAlaGluGlyGlyValValGluGlyThrGlyGluSerLeuProGluLys	432
DB	1486	-----AAACCGGAGGTGGAAAGAAACAGCAGGGTCTGTGCCAGCTGAAGAA	1530
QY	433	LeuAlaGluProGlnGluValProGlnGluAlaGluProAlaGluGluLeuMetLysSer	452
DB	1531	TTGGTTGAAATGGATCGAGAACCTCAGGAAGCGCAACCTGCCAAGAGAGTGGTGAAGCTC	1590
QY	453	ArgGluMetCysValSerGlyGlyAspHisThrGlnLeuThrAspLeuSerProGluGlu	472
DB	1591	AAAGAAACGTGTGTTTCCGAGAGGACCTACACAGGAGCTGACCTCAGCTCTGATGAG	1650
QY	473	LysThrLeuProLysHisProGluGlyIleValSerGluValGluMetLeuSerSerGln	492
DB	1651	AAGGTGCTGTCCAAACCCCGGAGGCGTGTGAGTGAGTGGAATGCTGTATCACAG	1710
QY	493	GluArgIleLysValGlnGlySerProLeuLysLysLeuPheSerSerSerGlyLeuLys	512
DB	1711	GAGAGAATGAAGTGCAGGGAAGTCCACTAAAGAAGCTTTTACCAGCAGCTGGCTTAAAA	1770
QY	513	LysLeuSerGlyLysLysGlnLysGlyLysArgGlyGlyGlyGluGluProGly	532
DB	1771	AAGCTTTCGGAAGAAACAGAAAGGGAAGAA---GGAGGAGGAGACCGAGGAATCAGGG	1827
QY	533	GluTyrGlnHisIleHisThrGluSerProGluSerAlaAspGluGlnLysGlyGluSer	552
DB	1828	GAGCACACTCAGGTTCAGCCGATTCCTCCGACAGCCAGGAGGAGCAAGGGCGGAGAGC	1887
QY	553	SerAlaSerSerProGluGluProGluGluThrThrCysLeuGluLysGlyProLeuGlu	572
DB	1888	TCTGCTCATCTCCCTCAGGAGCCGAGGAGATCACGTGTCTGGAAGAGGCTTTAGCCGAG	1947
QY	573	AlaProGluAspGlyGluAlaGluGlyThrThrSerAspGlyGluLysLysArgGlu	592
DB	1948	GTGACAGCAGATGGGAAGCTGAAGAAAGGAGGCTACTTCCGATGGAGAGAAAGAGAA	2007
QY	593	GlyIleThrProTrpAlaSerPheLysLysMetValThrProLysLysArgValArgArg	612
DB	2008	GGTGTCACTCCCTGGGCTTCATTCAAAAGATGGTGCAGCCCAAGAGAGCGGTGTAGACGG	2067
QY	613	ProSerGluSerAspLysGluGluGluLeuGluLysValLysSerAlaThrLeuSerSer	632
DB	2068	CCTTGGAAGATGATAAGAAAGATGAGCTGGCAAGGTCAAGAGCGCTACTCTTGTCTTCC	2127
QY	633	ThrAspSerThrValSerGluMetGlnAspGluValLysThrValGlyGluGluGlnLys	652
DB	2128	ACCGAGAGCAGCAGCTCTGAAATGCAAGAAAGAAATGAAAGGGAGCGGTGGGAAGACCAAG	2187

Db 4321 ACAGTGAATGCCCCATCATAGATGGGCAAGGAAGTACAGCTTTGGAGGAGCCCT 4380
 QY 1353 ProSerLeuProAspGlnAspLysAlaGlyCysIleGluValGlnValGlnSerLeuAsp 1372
 Db 4381 CCTCCCTCGCTAGTCAAGAGGAGGAGTATGACCAAAAATCAAGTTCAAGCTCGAG 4440
 QY 1373 ThrThrValThrGlnThrAlaGluAlaValGluLysValIle-----GluThrValVal 1390
 Db 4441 GCATCATCTCTTAACAGCGGCTGCAGAGGAGGAAGGTCTTAGAGAAATCGCCAAC 4500
 QY 1391 IleSerGluThrGlyGluSerProGluCysValGlyAlaHisLeuLeuProAlaGluLys 1410
 Db 4501 ATTTAGAAACAGGTGAACGTTGGACCTGCAGCTGCACATTTAGTTCTGGAGAGAAA 4560
 QY 1411 SerSerAlaThrGlyGlyHisTrpThrLeuGlnHisAlaGluAspThrValProLeuGly 1430
 Db 4561 TCCTCTGAAAAAATGAAGACTTTGCCGCTCATCCAGGGGAAGATGCTGTGCCACAGGG 4620
 QY 1431 ProGluSerGlnAlaGluSerIleProIleIleValThrProAlaProGluSerThrLeu 1450
 Db 4621 CCCGACTGTGAGGCAAAATCGACCAAGTATAGTATCTGCTACTACCAAGAAAGGCTTA 4680
 QY 1451 HisProAspLeuGlnGlyIleSerAlaSerGlnArgGluArgSerGluGluGluAsp 1470
 Db 4681 AGTTCCGACCTGGAGAGAGAGAAAACCATCTCACTGAAGTGGAGTCAATGATCGAT 4740
 QY 1471 LysProAspAlaGlyProAspAlaAspGlyLysGluSerThrAlaIleGluLysValLeu 1490
 Db 4741 GAGCAGGTGCTGCCAGGAGGTC-----AAAGTGAGTGTAGCAATTGAGGAGGATTA 4794
 QY 1491 LysAlaGluProGluIleLeuGluLeuGluSerLysSerLysIleValLeuAsnVal 1510
 Db 4795 GAGCCTGAAAATGGATTTTGGAACTTGAGACCAAAAGCAGTAACTGTGCCAAACATC 4854
 QY 1511 IleGlnThrAlaValAspGlnPheAlaArgThr---GluThrAlaProGluThrHisAla 1529
 Db 4855 ATCCAGACACCGTTGACCAAGTTGTAGTACAGAAAGAACAGCCACCAATATTGTAAG 4914
 QY 1530 TyrAspSerGlnThrGlnValProAlaCysArgLeuAspSerArgGluProAsnArgCys 1549
 Db 4915 TCTGAGTTACAGACACAGCTCAGCTGATAAAGCTCAGAGCCAGGACGCTGGACAGAA 4974
 QY 1550 TrpThrLys-----MetLysAspAla-LysMetLysHisProValProGlu 1564
 Db 4975 ACGAGAGAGAGGAGAGGAGAACCTCAGCGCTCTGCACAGGATGAACCAATATTCTCA 5034
 QY 1564 nProArgGlu-----AspLeuGlnValLeuThrValLeuGluAl 1577
 Db 5035 GCCAAGAGGAGTCAAGTCAACCGCAGTGGCAACAGCACATCTGTATATTTCCAAAGAC 5094
 QY 1577 aTrpAlaGlnProArgLysCysLeuProArgLeuGlnLeuLysAlaProVal 1594
 Db 5095 ATGAGTGAAGCCTCAGAAAAGACCATCATCTGTTGAGGTAGAGGTTCCACTG 5146

RESULT 14
 US-60-245-228-517/c
 ; Sequence 517, Application US/60245228
 ; GENERAL INFORMATION:
 ; APPLICANT: Beasley, Ellen
 ; TITLE OF INVENTION: ISOLATED HUMAN CYCLASE PROTEINS, NUCLEIC
 ; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN CYCLASE PROTEINS, AND USES
 ; TITLE OF INVENTION: THEREOF
 ; FILE REFERENCE: CL000878
 ; CURRENT APPLICATION NUMBER: US/60/245,228
 ; CURRENT FILING DATE: 2000-11-03
 ; NUMBER OF SEQ ID NOS: 630
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 517
 ; LENGTH: 6607
 ; TYPE: DNA
 ; ORGANISM: HUMAN
 ; US-60-245-228-517

Alignment Scores: 3.7e-213 Length: 6607
 Pred. No.: 4325.50 Matches: 969
 Percent Similarity: 68.83% Conservative: 188
 Best Local Similarity: 57.64% Mismatches: 408
 Query Match: 53.58% Indels: 116
 DB: 68 Gaps: 33
 US-09-902-432-4 (1-1596) x US-60-245-228-517 (1-6607)
 QY 1 MetGlyValAlaGlySerSerThrGluGlnArgSerProGluGln---ProAlaGlySerAsp 19
 Db 6413 ATGGGCGCGGGAGCTCCACCGAGCAGCGAGCCCGGAGCAGCGCCCGGAGGAGCTCC 6354
 QY 20 ThrProSerGluLeuValLeuSerGlyHisGlyProAlaAlaGluAlaSerGlyAlaAla 39
 Db 6353 ACGCGGCTGAGCGCGAGCCAGCGCGCGCGCCCTGGGCGAG-----GCGGCG 6303
 QY 40 GlyAspProAlaAspAla-----AspProAlaThrLysLeuProGln 53
 Db 6302 CCAGACCCCGCGCGCCCAACAATCGCTGCTGGGACCCCGCCCAAGCTCTTACAG 6243
 QY 54 LysAsnGlyGlnLeuSerSerValAsnGlyValAlaGluGlnGlyAspValHisValGln 73
 Db 6242 AAGAATGGTCAGCTGTCCACCATCAATGGCTAGCTGAGCAAGATGAGCTCAGCCCTCCAG 6183
 QY 74 GluLysAsnGlnGluGlyGln-----80
 Db 6182 GAGGCTGACCTAAATGCCAGAAAGGAGCCCTGAACGGTCAAGGAGCCCTAAACAGCCAG 6123
 QY 81 ---GluGluGluValValAspGluAspValGlyGlnArgGluSerGluAspValArgGlu 99
 Db 6122 GAGGAAGAAGAAGTCAATTGTCAAGGAGTTGGACAGAGAGACTCTGAAGATGTGAGCGAA 6063
 QY 100 LysAspArgValGluGluMetAlaAlaAsnSerThrAlaValGluAspIleThrLysAsp 119
 Db 6062 AGAGACTCCGATAAAGAGATGGCTACTAAGTCAAGCGGTGTTTCAACGACATCAAGATGAT 6003
 QY 120 GlyGlnGluGluThrSerGluIleIleGluGlnIleProAlaSerGluAsnValGlu 139
 Db 6002 GGCAGGAGGAGACACCCGAAATAATCGAACAGATTCCTTCTTCAGAAAGCAATTTAGAA 5943
 QY 140 GluMetValGlnProAlaGluSerGlnAlaAsnAspValGlyPheLysLysValPheLys 159
 Db 5942 GAGCTAACACACACCACTGAGTCCAGGCTAATGATATTGGATTAAAGAGGTGTTTAAAG 5883
 QY 160 PheValGlyPheLysPheThrValLysLysAspLysAsnGluLysSerAspThrValGln 179
 Db 5882 TTTGTTGGCTTTTAAATTCACGTGAAAAAGGATTAAGACAGAGAGAGCCTGACACTGTCCAG 5823
 QY 180 LeuLeuThrValLysLysAspGluGlyGluValAlaGluAlaSerValGlyAlaGlyAsp 199
 Db 5822 CTACTACTGTGAAGAAGATGAGGGGAGGAGCA-----GCAGGGGCTGGCGAC 5772
 QY 200 HisGlnGluProSerValGluThrAlaValGlyGluSerAlaSerLysGluSerGluLeu 219
 Db 5771 CACCAGGACCCAGCCT-----GGGGCTGGAGAGAGCAGCATCCAAAGAGAGCAACCC 5718
 QY 220 LysGlnSerThrGluLysGlnGluGlyThrLeuLysGlnGlnSerSerThrGluIle 239
 Db 5717 AAACAATCTACAGAGAAACCCGAAAGAGACCTTGAACCGTGAACAAAGCCACGAGAAAT 5658
 QY 240 ProLeuGlnAlaGluSerAspGlnAlaAlaGluGluAlaLysAspGluGlyGluGlu 259
 Db 5657 TCTCCCCAGCCGAATCTGGCCNAGCAGTG---GAGAAATGCAAGAGAGAGAGAGAGAG 5601
 QY 260 LysGlnGluLysGluProThrLysSerProGluSerProSerSerProValAsnSerGlu 279
 Db 5600 AAACAAGAAAAAGAACCTAGCAAGTCTGCAGAAATCTCCGACTAGTCCCGTGCACAGTGAA 5541
 QY 280 ThrThrSerSerPheLysLysPheThrHisGlyTrpAlaGlyTrpArgLysLysThr 299
 Db 5540 ACAGGATCAACCTTCAAAAAAATTTCTTCACTCAAGGTTGGCGCGCTGGCGCAAAAAGACC 5481

QY	300	SetPhelysIySerylgIuAspAspIeuGIuThAlaGIuIyArGlyeGIuGIu	319
Db	5480	AGTTTCAGGAAGCCGAAAGGAGGATGAACTGGAACTTCACAGAAAGAAAGGAAACAAAG	5421
QY	320	AlaGIuIySValAspGIuGIuGIuIySgIuIyThGluProAlaSerGIuGIu-----	337
Db	5420	CCAGAAAGATGACACAGAAAGAGACGGAAGGCGAGAGTTGCTCCGAGAAACTGACC	5361
QY	338	-----GInGIuProAlaGIuAspThrAspGIuAlaIyIeuSer	350
Db	5360	GCCTCCAGCAAGCCACCAACAGAGCCGGCAGAAAGTCCCAAGAGCCCGATTATCA	5301
QY	331	AlaAspIyTgIuIySValGIuIeuProIeuGIuAspGIuAlaGIyAspIeuGIuAlaSer	370
Db	5300	GCTGAATATGAGAAAGTTGAGCTGCTCCACAGAGCAAGTCAATGAGAAATGAAATCCAC	5241
QY	371	SerGIuGIuIySgIuAlaProIeuAlaThGIuIuAlPhaAspGIuIySmerGIuIuAlaIis	390
Db	5240	TCTGAAGGAAACCTGCTCGTGGCAGCAAGATTTGATGAGAAATGAAATCCAC	5181
QY	391	GIn---GIuValAlaIaGIuIuAlaIisValSerThrValGIuIySgIuGIuGIuGIu	409
Db	5180	CAAGAAAGGCTGTGGCCGGAAGTCCAGCTGAGCAACCTGAGAGAGAAACGAAAGAG	5121
QY	410	GIyGIyGIyGIuIuAlaGIuGIyValaIyValaGIuGIyThGIyGIuIeuPro	429
Db	5120	-----AAACCGAGGTGGAAGAAACAGCAGGGCTGTGCCA	5085
QY	430	ProGIuIySLeuAlaGIuProGIuGIuIuAlProGIuIuAlaGIuProAlaGIuGIuIeu	449
Db	5084	GCTGAAGATTGTGGAAGTGAATGATGCAGAACTCAGAGAAAGCCGACCTGCAGAGAGCTG	5025
QY	450	MetIySerAspGIuIeuCysValSerGIyGIyAspHisThrGIuIeuThrAspIeuSer	469
Db	5024	GTGAAGCTCAAGAAAGAGCTGTGTTCCGGAGAGAGACCTTACACAGGAGCTGACTCAT	4955
QY	470	ProGIuGIuIySThrIeuProIySHisProGIuGIyIleValSerGIuIuAlaGIuIeu	489
Db	4964	CCTGATGAGAAAGTGCTGTCCAAACCCCGAGAGCGTTGTGATGAGGTGAATGCTG	4905
QY	490	SerSerGIuGIuIuGIyIleIySValGIuGIySerProIeuIyIyIeuIeuPheserSerSer	509
Db	4904	TCATACACAGGAGAAATGAAGTGCAGAGGAAGCTCACTAAAGAGCTTTTATCCAGGACT	4845
QY	510	GIyIeuIySLeuSerGIyIySgIuIySgIuIyAspGIyGIyGIyAspGIu	529
Db	4844	GCCTTAAAAAGCTTTCTCGAAGAAACAGAAAGGAAAAAGA---GGAGAGAGAGACAG	4788
QY	530	GIuProGIyGIuIyTgIuIuAlaIisHisThrIuSerProGIuSerAlaAspGIuIuIyS	549
Db	4787	GAATCAGGGGAGCACACACAGATTCCAGCCGATTCTCCGACAGCCAGAGAGGCAAAAG	4728
QY	550	GIyGIuIeuSerAlaSerSerProGIuGIuIuProGIuGIuIuThrThrCysIeuGIuIySgIy	569
Db	4727	GCGCAGACCTCGCTCATCCCTCGAGAGACCCGAGAGAAATCAACGTGTCTGGAAGAGGC	4668
QY	570	ProIeuGIuAlaProGIuIuAspGIyGIuIuAlaGIuGIuGIyThrThrSerAspGIyGIuIyS	589
Db	4667	TTAGCCGAGGTGCAGACAGAGATGGGGAAGCTGGAAGAAAGAGACTTACTTCGATGGAAGAA	4608
QY	590	IySArGIuGIyIleIeuProIyAlaSerPheIyIySmerValIeuProIyIySArGIy	609
Db	4607	AAAAAGAAAGGTTCATCTCCCTGGGCTTACTTCAAAAAGATGTGACGCCCAAGAGAGCT	4588
QY	610	ValaIyArGIuProSerGIuIeuAspIySgIuGIuGIuIeuGIuIySValIySserAlaThr	629
Db	4547	GTTAAGCGCTTCGGAAGATGAATGAAGATGAGACTGGAACAAAGCTCAAGAGCTTACC	4488
QY	630	LeuSerSerThrAspSerThrValSerGIuIeuGIuAspGIuIuValIySerThrValGIyGIu	649
Db	4487	TTGTCTTCCACCGAGAGACAGCTCTGGAATCAAGAAAGAAATGAAGAGGAGCGTTGAA	4428

[illegible]

Db 3347 CAGAGGTGAGAGGTGCGTACCTGACATAGAACAGACGAGAGGCGGACTCAAGAGGTC 3288
QY 1022 LeuGlnAlaValAlaAspTysValLysGluGluSerGlnValProAlaThr 1038
Db 3287 CTCAGGACGTGGCAGAAAAGTGAAGAGAGAAATCCAGAGTCCCTGGCACCGGTGGGCCA 3228
QY 1039 -----GlnThrValGlnAlaThrGlySerLysAlaLeuGluLysValGluGlu 1054
Db 3227 GAAGATGTGCTTCAGCGCTGGCAGAGACGAC-----GAGGCGAAGAAGA 3186
QY 1055 ValGluGluAspSerGluValLeuAlaSerGluLysGluLysAspValMetProLysGly 1074
Db 3185 CAGAAGACGAGCTGAGCGTGGGTCTGAGAGAGAGACGATGATGTTGTTGAAGTA 3126
QY 1075 ProValGlnGluAlaGlyAlaGluHisLeuAlaGlnGlySerGluThrGlnAlaThr 1094
Db 3125 GATGCTCAGGAGGCAAAACCTGAGCCCTTTACACAGGGAAGGTGGTGGGCGACACC 3066
QY 1095 ProGluSerLeuGlu-----ValProGluValThrAlaAspValAspHis-----Val 1110
Db 3065 CCAGAAGGCTTTGAAAAGCTCCTCAAGTCACAGAGACGATGATGCTGAGCTTGTA 3006
QY 1111 AlaThrCysGlnVal-----IleLysLeuGlnGlnLeu-----MetGluGln 1124
Db 3005 ACCATTTGTCACGCCGAAACCTTAGCTGGGGTAAATTCACAGGAGATGGTATGGAACAG 2946
QY 1125 AlaValAlaProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrPro 1144
Db 2945 GCTATCCCCCTCAGTCCGTTGGTGAACCCCTACAGACAGTGAAGTATGATGGAACACCCC 2886
QY 1145 LeuAlaAspSerAspThrAlaAspGlyThrGlnGlnAspGluThrLeuAspSerGlnAsp 1164
Db 2885 GTAGCGGACTTTGACGCCAGCAGCAACCCAGAAAAGCAGATTTGTGGAATCCATGAG 2826
QY 1165 SerLysAlaThrAlaAlaValAlaGlnSerGlnValThrGluGluAlaAlaThrAla 1184
Db 2825 GAGAAATGAGTCCGATCTGTGATCCAGTACGGGGGACAGACAGAGGAGGAGTTCCTGCA 2766
QY 1185 GlnLysGluGluProSerThrLeuProAsnAsnValProAlaGlnGluGluHisGlyGlu 1204
Db 2765 CAGAAAGAGAGGCTCCAGCA---CCTTCAGTTTGTGTTCCAGGAGAACTAAAGAA 2709
QY 1205 GluProGly---ArgAspValLeuGluProThrGlnGlnGluLeuThrAlaAlaVal 1223
Db 2708 CAATCAAGATGGAAGACACTCTAGAGCATACAGATAAAGAGGTGTCAGTGGAACTGTA 2649
QY 1224 ProValLeuAlaLysThrGluValGlyGlnGluGlyGluValAspThrLysAspGlyGlu 1243
Db 2648 TCCATTCTGTCAAGAGCTGAGGGGACTCAA-----GAGGCTACCATGATGCTGATGAG 2595
QY 1244 LysValLys-----GluGluGlnGluValPheValHisSerGly----- 1256
Db 2594 AAACCAAGACAGCTACCATTTTTCGAGAGGACTTGAGGGGTCTATAGACACAGGCATAACA 2535
QY 1257 ProAsnSerGlnLysAlaAlaAspValThrTyrAspSerGluValMetGlyValAlaGly 1276
Db 2534 GTCAGTCGGGAAAAGGTCAGTGAAGTTGCCCTTAAAGGTGAAGGAGCAGAAAGCTGAA 2475
QY 1277 CysGlnGluLysGluSerThrGluValGlnSer-----LeuSerLeuGlu 1291
Db 2474 TGTAAGAGATGATGCTCTTGAACCTGCAGAGTCACGCTAAGTCTCCTCCATCCCGGTG 2415
QY 1292 GluGlyGluMetGluThrAspValGluLysGluLysArgGluThrLysProGluGlnVal 1311
Db 2414 GAGAGAGATGTTAGTTCAAGTCGAAGGAGAGAAACAGACAGACGCCAACCCATGTG 2355
QY 1312 SerGluGluGly---GluGlnGluThrAlaAlaProGluHisGluGlyThrTyrGlyLys 1330
Db 2354 AATGAAGAGAAGCTTGAGCACGCAAAACAGCTGTTACCGTATCTCAAGAGGTGCTAAGCAG 2295
QY 1331 ProValLeuThrLeuAspMetProSerSerGluArgGlyLysAlaLeuGlySerLeuGly 1350
Db 1350

Db 2294 CTCCTCCACACAGTGAATGTGCCCATCATAGATGGGCAAGAAAGTACAGAGTTTGGAA 2235
QY 1351 GlySer---ProSerLeuProAspGlnAspLysAlaGlyCysIleGluValGlnValGln 1369
Db 2234 GGAAGCCCTCCTCCCTCCCTAGTCAAGAGGAGGAGTATGACCAAAAAATCAAGTTTCA 2175
QY 1370 SerLeuAspThrThrValThrGlnThrAlaGluAlaValGluLysValIle-----Glu 1387
Db 2174 AGCTCTGAGGCATCATCTTAACAGCGGTGCAGAGGAGGAAAGGCTCTTAGGAGAA 2115
QY 1388 ThrValValIleSerGluThrGlyGluSerProGluCysValGlyAlaHisLeuLeuPro 1407
Db 2114 ACTGCCAACATTTTAGAACACAGTGAACCTTGGAGCTGCAGGTCACATTTAGTTCTG 2055
QY 1408 AlaGluLysSerSerAlaThrGlyGlyHisTrpThrLeuGlnHisAlaGluAspThrVal 1427
Db 2054 GAAGAGAAATCCTCTGAAAAAATGAAGACTTTGCCGCTCATCCAGGGGAAGATGCTGTG 1995
QY 1428 ProLeuGlyProGluSerGlnAlaGluSerIleProIleIleValThrProAlaProGlu 1447
Db 1994 CCCACAGGGCCCGACTGTGAGGCAAAATCGACACAGTGTATGATCTGCTACTACCAAG 1935
QY 1448 SerThrLeuHisProAspLeuGlnGlyIleSerAlaSerGlnArgGluArgSerGlu 1467
Db 1934 AAAGGCTTAAGTTCGACCTGGAAGAGAGAAACACATCACTGAAGTGGAGTCAGAT 1875
QY 1468 GluGluAspLysProAspAlaGlyProAspAlaAspGlyLysGluSerThrAlaIleGlu 1487
Db 1874 GAAGTCGATGACGAGGTTCTTCCAGGAGGTC-----AAAGTGAAGTGTAGCAATTGAG 1821
QY 1488 LysValLeuLysAlaGluProGluLysLeuGluLeuGluSerLysSerAsnLysIleVal 1507
Db 1820 GAGGATTTAGAGCCCTGAAATGGGATTTTGGAACTTGAGACCAAAAGCAGTAACTGTG 1761
QY 1508 LeuAsnValIleGluThrAlaValAlaAspGlnPheAlaAspThr---GluThrAlaProGlu 1526
Db 1760 CAAAACATCATCCAGACGCGTTGACCGAGTTGTACGTACAGAAAGAACAGCCCGAA 1701
QY 1527 ThrHisAlaTyrAspSerGlnThrGlnValProAlaCysArgLeuAspSerArgGluPro 1546
Db 1700 ATGTTGACGCTGAGTTTACAGACACAGCTCAGCTGATATAAAGCTGACAGCAGCGCT 1641
QY 1547 AsnArgCysTrpThrLys-----MetLysAspAla---LysMetLysHisPr 1561
Db 1640 GGACAGGAAACGGAGAAAGAGAGAGGAAACCTCAGCGCTCTGCACAGGATGAACACCA 1581
QY 1561 OValProGlnProArgGlu-----AspLeuGlnValLeuThrVa 1574
Db 1580 ATTACTTCAGCCAAAGAGGAGGAGTCAGAGTCAACCGCAGTGGGACACGACATTTGTAT 1521
QY 1574 LLeuGluAlaTrpAlaGlnProArgLysCysLeuProArgLeuGlnLeuLysAlaProVa 1594
Db 1520 TCCAAAGACATAGTGAAGCCTCAGAAAGACCATGACTGTTGAGGTAGAGTCCACT 1461
QY 1594 I 1594
Db 1460 G 1460

RESULT 15

US-09-644-867-6364
; Sequence 6364, Application US/09644867
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Donovan, Michael J.
; APPLICANT: Leiby, Kevin R.
; APPLICANT: Culpepper, Janice A.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.1173-001
; CURRENT APPLICATION NUMBER: US/09/644,867
; PRIOR FILING DATE: 2000-08-28
; PRIOR APPLICATION NUMBER: 60/151,061
; PRIOR FILING DATE: 1999-08-27

```

? NUMBER OF SEQ ID NOS: 8090
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 6364
? LENGTH: 8660
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: misc feature
? LOCATION: (1)_(8660)
? OTHER INFORMATION: n = A,T,C or G
? OS-09-644-867-6364

```

Alignment Scores:	
Pred. No.:	1,5e-212
Score:	4316.00
Percent Similarity:	68.898
Best Local Similarity:	57.638
Query Match:	53,468
AB:	25
	Gaps:
	34

US-09-902-432-4 (1-1596) X US-09-644-867-6364 (1-8660)

Oy	1	MetGlyAlaGlySerSerThrgluGlnAArgserProGluGln---	ProAlaGlySerAsp	19
Db	183	ATGGCGCGGGAGCTCCACGAGAGCGACGCCGGAGCACGCCCCGAGGGAGCTCC		242
Oy	20	ThrProSerGluLeuValIleuSerGlyYHtsglyProAlaIaGluIaSer---	GlyAla	38
Db	243	ACGCCGCTGAGCCCGAGCCACGAGGGGGGGGGCCCTCTGGCGGAGGGGGCGGACGAC		302
Oy	39	AlaGlyAspProAlaAspAla-----AspProAlaThrIlybLeuProGlnIlyAsnGly		56
Db	303	ACCGGGGACCCCGGCATCTGCTCGGACCCCGCCACCAAGCTCTCTACAGAAAGAAAGT		362
Oy	57	GluIleuSerSerValAsnGlyValAlaGluGlnGlyAspValHtIsvaGlnGluGluAsn		76
Db	363	CAGCTGTCCACCATCAATGAGGCTGAGCGAAGATGAGCTCAGCTTCCAGAGGGTGAC		422
Oy	77	GlnGlnGlyGln-----GlnGlu		82
Db	423	CTAATATGCCGAGAAAGAGCCCTGAACGCTCAAGAGCCCTTAACACGACGAGAGAA		482
Oy	83	GluValValaAspGluAspValaGlyGlnArgGluSerGluAspValaArgGluIlyAspArg		102
Db	483	GAACTCATTTGCACGAGAGTTGGACAGAGAGACTCTGAATGATGAGCGAAAGAGACTCC		542
Oy	103	ValGluGluMetAlaIaIaAsnSerThraIaValaGluAspIleThrIlybAspGlyGlnGlu		122
Db	543	GATTAAGAGATGGCTACTAATCTACGCGGGTTCACGACATCACAGATGATGGCGAGAG		602
Oy	123	GluThrSerGluIleIleGluGlnIleProAlaSerGluAsnAsnValaGluGluMetVal		142
Db	603	GAGAACCCGAAT--ATCGAACAGATTCCTCTTCAGAAACCAATTTAGAAAGACTACA		658
Oy	143	GlnProAlaGluSerGlnAlaIaAsnAspValaGlyPheIlyIlyValaPheIlybPheValaGly		162
Db	660	CAACCCACTGAGTCCCAAGCTCAATGATATGTGATTAAAGAGGTGTTAAGTTGTGGC		719
Oy	163	PheIlybPheThrValaIlybIlybAspIlybAsnGluIlybSerAspThrValaGlnIleuLeuThr		182
Db	720	TTTAAATTCACTGTGAAAAAGATTAAGCAAGAGAGCTGACATGTCCAGCTACTACT		779
Oy	183	ValIlybIlybAspGluGlyGluGlyValaGluIuIaSerValaGlyAlaGlyAspHtsgGlnGlu		202
Db	780	GTGAGAAAGATGAGAGGGGAGGAGCA-----GCAGGGGCTGGCGACACCAAGAC		830
Oy	203	ProSerValaGluThrAlaValaGlyGluSerAlaSerIlybGluSerGluLeuIlybGlnSer		222
Db	831	CCCAAGCTT-----GGGCGTGGAGAAACGACCATCTCAAGAAAGCAACCCAAACATCT		884
Oy	223	ThrGluIlybGlnGluGlyThrIleuIlybGlnGluGlnSerSerThrGlnIleProIleGln		242
Db	885	ACAGAGAAACCCGAAAGAACCTTGAGGTGAGCAAAAGCCACGCAAGAAATTTCTCCCCA		944

QY	243	AlaGluSerAspGlnAlaAlaGluGluGluAlaLysAspGluGluGluGluGluGlnGlu	262
Db	945	GCCAAATCTGGCCAAAGACGTG---GAGGAATCCAAAGAGAAAGGAGAAACAAAGAA	1001
QY	263	LysGluProThrLysSerProGluSerProSerProValAsnSerGluThrThrsSer	282
Db	1002	AAAGAACCTTACGAAGTCTGCAGAAATCTCCGACTAATGCCGTGACAGTGAACAGAAATCA	1061
QY	283	SerPheLysLysPhePheThrHisGlyTyrAlaGlyTyrPArgLysLysThrSerPheLys	302
Db	1062	ACCTTCAAAAATTCCTTCACTCAAGGTGGCGCGGTGGCGCAAAAACCAAGATTTCAGG	1121
QY	303	LysSerLysGluAspAspLeuGluThrAlaGluLysArgLysGluGlnGluAlaGluLys	322
Db	1122	AAGCCAGAGAGAGAGTGAAGGAGAACTTCAAGAAAGAAAGAAACAAAGACCAAGAAAA	1181
QY	323	ValAspGluGluGluLysGlyLysThrGluProAlaSerGluGlu-----	337
Db	1182	GTAGACACAGAAAGAACCGAAAGGACAGAGGTTCCTCCGAGAAATCTACCGCGCTCCGAG	1241
QY	338	-----GlnGluProAlaGluAspThrAspGlnAlaArgLeuSerAlaAspTyr	353
Db	1242	CAAGCCACCCACAGAGACGAGGACGAGAAATGCCCAAGACCCCGGTTTACGCTGAATAT	1301
QY	354	GluLysValGluLeuProLeuGluAspGlnValAlaGlyAspLeuGluAlaSerSerGluGlu	373
Db	1302	GAGAAATGTAGAGCTGCCCTCAGAGAGACAAATCAATGAGTCTGCACAGAACCTTGTGAAG	1361
QY	374	LysCysAlaProLeuAlaThrGluValPheAspGluLysMetGluAlaHisGln--Glu	392
Db	1362	AAACCTGCTCGGTGGGACAGAAAGTTTGATGAGAAATGAAGTCCACCAACAAAG	1421
QY	393	ValValAlaGluValHisValSerThrValGluLysThrGluGluGluGluGlyGly	412
Db	1422	GTTGTGGCCAAAGTCCACGTCACACCTCGTAGAGAGAAACCGAAGACAG-----	1472
QY	413	GlyGluAlaGluGlyGlyValValGluGlyThrGlyGluSerLeuProGluLys	432
Db	1473	-----AAAGCGAGGTGGAAGAAACAGCAGCGGTCTGTGCTCAGCTGAAGAA	1517
QY	433	LeuAlaGluProGlnGluValProGlnGluAlaGluProAlaGluGluLeuMetLysSer	452
Db	1518	TTGGTTGAATGATGATGCAGAACTCAGAAAGCCGAACCTGCGCAAGAGACTGTGTAAGCTC	1577
QY	453	ArgGluMetCysValSerGlyGlyAspHisThrGlnLeuThrAspLeuSerProGluGlu	472
Db	1578	AAAGAAACGTGTGTTTCCGAGAGACACCTACACAGGAGAGCTGACTGCTCTGATAG	1637
QY	473	LysThrLeuProLysHisProGluGlyIleValSerGluValGluMetLysSerGln	492
Db	1638	AAGGCTGCTCCAAACCCCGGAAGCGTGTGATGAGAGTGAATGCTGTCAATCACAG	1697
QY	493	GluThrGlyLeuValGlnGlySerProLeuLysLysLeuPheSerSerSerGlyLeuLys	512
Db	1698	GAGGAATGAAGGTCGAGGAGAAATCTCAATMAAGCTTTTACACAGACTGACTTAAAA	1757
QY	513	LysLeuSerGlyLysLysGlnLysGlyLysArgGlyGlyGlyLysAspGluGluProGly	532
Db	1758	AAGCTTCTGGAAGAAACAGAAAGGAGAAAG---GAGAGAGAGAGACGAGGAATCAAGG	1814
QY	533	GluTyrGlnHisIleHisThrGluSerProGluSerAlaAspGluGlnLysGlyLysSer	552
Db	1815	GAGCACTACAGTTCACAGCCGATCTCCGAGACAGCAAGAGACCAAAAGGCGCAAGAGC	1874
QY	553	SerAlaSerSerProGluGluProGluGluThrThrCysLeuGluLysGlyProLeuGluGlu	572
Db	1875	TCTGCTCATCTCCCTGAGAGAGCCGAGAGAGATCACTGTCTGTGAAAAGGCTTACCGAG	1934
QY	573	AlaProGlnAspGlyGluAlaGluGluGlyThrThrSerAspGlyGluLysLysArgGlu	592
Db	1935	GTCGACGACGAGTGTGGGAAGCTGAAGAAAGGCTACTCTCGATGAGAGAGAAAAAGAGAA	1994

Qy	593	Gly	Thr	Pro	Trp	Ala	Ser	Pro	Leu	Val	Thr	Pro	Leu	Val	Arg	612
Db	1995	GGT	CTC	ACT	CCCT	CGG	CAT	CAT	TTC	AAAA	AGAT	CGT	GAC	CCCA	GAAC	2054
Qy	613	Pro	Ser	Glu	Ser	Asp	Leu	Glu	Leu	Glu	Val	Leu	Ser	Val	Ser	632
Db	2055	CC	TT	CGG	AA	GTG	AT	GAT	AG	CT	TGG	CA	AG	GG	CT	2114
Qy	633	Thr	Asp	Ser	Thr	Val	Ser	Glu	Met	Gln	Asp	Glu	Val	Thr	Val	652
Db	2115	ACC	GAG	AG	CAC	AG	CGT	CT	GAA	AT	GCA	AGA	AAAT	GAA	AG	2174
Qy	653	Pro	Glu	Glu	Pro	Leu	Arg	Val	Asp	Thr	Ser	Val	Ser	Trp	Glu	672
Db	2175	CCG	GA	GA	CA	CA	AG	CG	CA	AG	TGG	GAT	TAC	CT	TGG	2234
Qy	673	Gly	Ser	Ser	Leu	Ser	Val	Ser	Arg	Ala	Arg	Leu	Ala	Ser	Leu	692
Db	2235	GG	AT	CAT	CCA	GA	AA	GAG	CA	AGG	AGG	TCT	CTT	CTG	ATG	2294
Qy	693	Thr	Leu	Gly	Gly	Asp	Ser	His	Arg	Lag	Glu	Ala	Ser	Leu	Ala	712
Db	2295	GCA	AT	G	GG	AG	GAC	CAC	CA	AA	AG	CTG	AT	G	GG	2354
Qy	713	Asp	Ala	Val	Pro	Leu	Ser	Thr	Gln	Gln	Asp	Gln	Ala	Gln	Ser	732
Db	2355	GAC	GG	AT	CTT	CGT	T	CCCA	AA	CA	AT	GAT	CAG	GG	CA	2414
Qy	733	Pro	Ala	Gly	Ser	Pro	Ser	Glu	Gly	Glu	Gly	Val	Ser	Thr	Trp	752
Db	2415	CA	AG	CT	GGA	AG	CC	CT	ACC	GA	GGG	AGG	CG	CTT	CC	2474
Qy	753	Val	Thr	Pro	Arg	Leu	Ser	Ser	Leu	Ser	Leu	Glu	Leu	Ala	Glu	769
Db	2475	GT	CA	GC	CA	GA	AAAA	AAAT	CA	AA	GT	CA	AG	CT	GGA	2534
Qy	770	Ser	Val	Glu	Glu	Leu	Ser	Thr	Glu	Leu	Pro	Ser	Arg	Glu	Ser	788
Db	2535	GG	GT	CT	GGT	CT	AG	CA	ATT	CC	ACT	CC	CA	CA	CT	2594
Qy	789	Ser	Leu	Ser	Leu	Ser	Thr	Arg	Val	Ser	Thr	Arg	Val	Ser	Thr	808
Db	2595	TCA	AT	CA	AG	AA	GT	TTA	TTC	CTG	CA	GA	AG	AA	AG	2654
Qy	809	Ala	Thr	Val	Glu	Asp	Ser	Gly	Pro	Val	Glu	Leu	Asn	Glu	Asp	828
Db	2655	GCC	CT	GTT	GGA	CG	CG	CC	CA	AG	GG	CC	CA	CA		

Db	3075	 ::: 	GTTAGTGAGCGGAATTTGACCCCGGAAGCTGTGCACAGCTGCAGAAACTGCAGAGGCCAATG	3134
Qy	965	 ::: 	ArgThrGluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSerAla	984
Db	3135	 ::: 	GGTCCCGAAGAGAACCGAGCACTGCTGCTGAAGAGACACACAGAAATGGTGTGACGA	3194
Qy	985	 ::: 	ValSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGluVal	1004
Db	3195	 ::: 	GTCTCCCAAGTTAACCGACTCCCGACACACACAGAGGAGGCCACTCCGGTGCAGAGGTG	3254
Qy	1005	 ::: 	GluSerGlyValLeuAspThrGluGluGluArgGlnThrGlnAlaIleLeuGlnAla	1024
Db	3255	 ::: 	GAAGGTGGCTACTCCTCACATAGAAGACCAAGAGCGCGACTCAAGAGGTCTCTCCAGCA	3314
Qy	1025	 ::: 	ValAlaAspLysValLysGluGluSerGlnValProAlaThr-----	1038
Db	3315	 ::: 	GTGGCAGAAAAGTGAAGAGGATCCCGACTGCTGGCACCGGTGGCCAGAGAATGTG	3374
Qy	1039	 ::: 	---GlnThrValGlnArgThrGlySerLysAlaLeuGluLysValGluGluValGluGlu	1057
Db	3375	 ::: 	CTTCAGCCTGTGCAGAGACA-----GAGGCAGAAAGACCAAGAAGAG	3416
Qy	1058	 ::: 	AspSerGluValLeuAlaSerGlnLysGluLysAspValMetProLysGlyProValGln	1077
Db	3417	 ::: 	CAGGCTCAAGCGTGGGTCTCGAAGAAAGACGCGATAGTGTGTGAATAGTAGTCTCAG	3476
Qy	1078	 ::: 	GluAlaGlyAlaGluHisLeuAlaGlnGlySerGluThrGlyGlnAlaThrProGluSer	1097
Db	3477	 ::: 	GAGGCAGAAAACCTGAGCCTTTTACACAAAGGAAGTGTGGGGCGACACCACCCAGAAAGC	3536
Qy	1098	 ::: 	LeuGlu---ValProGluValThrAlaAspValAspHis-----ValAlaThrCys	1113
Db	3537	 ::: 	TTTGAAAAAGCTCCTCAAGTCCACAGAGACATAGAGTCCAGTGCAGCTGTGTAACCACTGT	3596
Qy	1114	 ::: 	GlnVal-----IleLysLeuGlnGlnLeu---MetGluGlnAlaValAla	1127
Db	3597	 ::: 	CAAGCCGAAAACCTTAGCTGGGTAAATAACACAGAGAGATGGTGCATGGAAACAGGCTATCCCC	3656
Qy	1128	 ::: 	ProGluSerSerGlnThrLeuThrAspSerGluThrAsnGlySerThrProLeuAlaAsp	1147
Db	3657	 ::: 	CCTCACTCGGTGAACCCCTCACAGACATGAGACTGATGTAAGAACGCCCCGTCAGCCGAC	3716
Qy	1148	 ::: 	SerAspThrAlaAspGlyThrGlnGlnAspGluThrIleAspSerGlnAspSerLysAla	1167
Db	3717	 ::: 	TTTGACGCCACCGGCACACCCAGAAAGACGAGATGTGGAAATCCATGAGGAGAAATCAG	3776
Qy	1168	 ::: 	ThrAlaAlaValArgGlnSerGlnValThrGluGluGluAlaAlaThrAlaGlnLysGlu	1187
Db	3777	 ::: 	GTCGCATCTGTTACCCAGTCAGGGGCACACAAAGACAGAGCGAGCTTCTCTGCACAGAAAGAG	3836
Qy	1188	 ::: 	GluProSerThrLeuProAsnAsnValProAlaGlnGluGluHisGlyGluGluProGly	1207
Db	3837	 ::: 	AGGCCTCCAGCA--CCTTCCAGTTTGTGTTCCAGGAAGAAACTAAAGAAACAATCAAAAG	3893
Qy	1208	 ::: 	---ArgAspValLeuGluProThrGlnGlnLeuThrAlaAlaAlaValProValLeu	1226
Db	3894	 ::: 	ATGAAGACACTCTAGAGCATACAGATAAAGAGGTGTCAGTGGAAACTGTATTCATCTCTG	3953
Qy	1227	 ::: 	AlaLysThrGluValGlyGlnGluGlyGluValAspTrrLeuAspGlyGluLysValLys	1246
Db	3954	 ::: 	TCAAAGACTCAGGGGACTCAA-----GAGGCTGACCATGCTGTGATGAGAAAAACCAA	4007
Qy	1247	 ::: 	-----GluGluGlnGluValPheValHisSerGly-----ProAsnSer	1259
Db	4008	 ::: 	GACGTACCATTTTTTCAAGGACTCTGAGGGGTCTATATAGACACAGGCATTAACAGTCCGG	4067
Qy	1260	 ::: 	GlnLysAlaAlaAspValThrTyraSerSerGluValMetGlyValAlaGlyCysGlnGlu	1279
Db	4068	 ::: 	GAAGAAGTCACTGAAGTTGCCCTTTAAGGTGAAGGCACAGAAGAACTGAATGTATAAAG	4127
Qy	1280	 ::: 	LysGluSerThrGluValGlnSer-----LeuSerLeuGluGluGlyGlu	1294

Db 4128 GATGATGCTCTTGAAGTCAGAGTCAGCTAAGTCTCCATCCCGGAGAGAGAG 4187
QY 1295 MetGluThrAspValGluIuysGluIuysArgGluThrLysProGluIuValSerGluIu 1314
Db 4188 ATGATGATTCAGAGTCGAAAGGAGGAAACAGAGCAGAGCAACCATGTGATGAAAGAG 4247
QY 1315 Gly---GluGluGluThrAlaIaProIuHisGluGlyThrTyGlyLysProValLeu 1333
Db 4248 AAGCTTGACAGCAAGAACAGCTGTACCTGATTCGAAAGAGTCAGTAAGAGCTCTCCAG 4307
QY 1334 ThrLeuAspMetProSerSerGluArgGlyLysAlaLeuGlySerLeuGlyLysSer--- 1352
Db 4308 ACAGTCATGATGCTCCATCATATGATGGGCAAGAGAGTTCAGCATTTGGAGAGAAAGCCCT 4367
QY 1353 ProSerLeuProAspGlnAspLysAlaGlyCysIleGluValGlnValGlnSerLeuAsp 1372
Db 4368 CCTCCCTGCTGAGTCAGAGAGGAGGAGATGACCAAAATTCAGATTCAGAGCTCTGAG 4427
QY 1373 ThrThrValThrGlnThrAlaGluIaValGluLysValIle-----GluThrValVal 1390
Db 4428 GCATCATTCACCTTACACGCTGCTGAGAGAGAAAGGCTTAGAGAGAACTGCCAAC 4487
QY 1391 IleSerGluThrGlyGluSerProGluCysValGlyAlaHisLeuLeuProAlaGluLys 1410
Db 4488 ATTTTAGAAACAGGTGAAACGTTGGAGCCTGCGAGCTGCACATTAGTTCTGGAAGAA 4547
QY 1411 SerSerAlaThrGlyGlyHisIleThrLeuGlnHisAlaGluAspThrValProLeuGly 1430
Db 4548 TCCTCTGAAATAATGATAGACTTTCCTCATCCAGGGAAGATGCTGTGCCACAGAG 4607
QY 1431 ProGluSerGlnAlaGluSerIleProIleIleValThrProAlaProGluSerThrLeu 1450
Db 4608 CCCGACTGTACGCAAAATCGACACAGATGATGATCTACTACTACCAAGAAAGGCTTA 4667
QY 1451 HisProAspLeuGlnGlyIuIleSerAlaSerGlnArgGluArgSerGluGluAsp 1470
Db 4668 AGTTCGACCTGAGAGAGAGAAACCATCATCTGAAAGTGAAGTCAGATGATCGAT 4727
QY 1471 LysProAspAlaGlyProAspAlaAspGlyLysGluSerThrAlaIleGluLysValLeu 1490
Db 4728 GAGCAGGTGCTTGCAGAGAGGTC-----AAGGTAGGTAGCAATTGAGAGGATTTA 4781
QY 1491 LysAlaGluProGluIleLeuGluIuSerLysSerAsnLysIleValLeuAsnVal 1510
Db 4782 GAGCCTGAAATGGGATTTGGAACTTGAGACCAAAAGCAGTAACCTTGCCAAAACATC 4841
QY 1511 IleGlnThrAlaValAspGlnPheAlaArgThr---GluThrAlaProGluThrHisAla 1529
Db 4842 ATCCAGACAGCCGTTGACGTTTGTACGTACAGAAAGAACAGCCAGAAATGTTGACG 4901
QY 1530 TyrAspSerGlnThrGlnValProAlaCysArgLeuAspSerArgGluProAsnArgCys 1549
Db 4902 TCTGAGTTACAGACACAGCTCAGTATAAAGCTGACAGCAGCAGCGTGGACAGAA 4961
QY 1550 TrpThrLys-----MetLysAspAla-LysMetLysHisProValProGlu 1564
Db 4962 ACGAGAAAGAGAGAGAGACCTCAGCGCTCTGACAGATGAAACACCAATTACTTCA 5021
QY 1564 nProArgGlu-----AspLeuGlnValLeuThrValLeuIuVal 1577
Db 5022 GCCAAAGAGAGAGTCAGAGTCACCGCAGTGGACAAAGCACATTCTGATATTTCCAAAGAC 5081
QY 1577 artpAlaGlnProArgLysCysLeuProArgLeuGlnLeuLysAlaProVal 1594
Db 5082 ATGAGTGAAGCTCAGAAAGAACATGACTGTTGAGGTAGAAAGTTCACACTG 5133

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 13, 2002, 00:45:36 ; Search time 1677 Seconds
(without alignments)
692.852 Million cell updates/sec

Title: US-09-902-432-4
Perfect score: 8073
Sequence: 1 WGAGSSTQRSSEQAGSDT.....ANAPRKCLPRLQKAPVSK 1596

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 422006 seqs, 364007625 residues

Total number of hits satisfying chosen parameters: 844012

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgn2_1/USPTO spool/US0902432/runat_12122002_141453_25888/app query.fasta_1.1735
-DB=Pending Patents NA_New -QFMT=fastap -SUFFIX=xnnp -MINMATCH=0 -1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -LOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptt -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0902432 @C@N 1 1 156 @runat 12122002 141453 25888 -NCPU=6 -ICPU=3
-NO XLPXY -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Pending Patents NA_New :

1: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq.*
2: /cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq.*
3: /cgn2_6/ptodata/1/pna/US07_NEW_COMB.seq.*
4: /cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq.*
5: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq.*
6: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq.*
7: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8030	99.5	5236	6 US-10-152-319A-2063	Sequence 2063, Ap
2	4341.5	53.8	5931	5 US-09-724-676-12735	Sequence 12735, A
3	4341.5	53.8	5931	5 US-09-724-676A-12735	Sequence 12735, A
4	4300.5	53.3	6608	1 PCT-US02-04915-43	Sequence 43, Appl
5	4104.5	50.8	5773	5 US-09-724-676-12736	Sequence 12736, A
6	4104.5	50.8	5773	5 US-09-724-676A-12736	Sequence 12736, A
7	492	6.1	9161	6 US-10-133-937-37	Sequence 37, Appl
8	431.5	5.3	7997	6 US-10-203-138A-6112	Sequence 6112, Ap
9	427	5.3	15231	6 US-10-152-319A-1780	Sequence 1780, Ap
10	418	5.2	14335	5 US-09-724-676-25032	Sequence 25032, A
11	418	5.2	14335	5 US-09-724-676A-25032	Sequence 25032, A

12	415	5.1	8161	5 US-09-724-676-7257	Sequence 7257, Ap
13	415	5.1	8161	5 US-09-724-676A-7257	Sequence 7257, Ap
14	415	5.1	8189	5 US-09-724-676-7240	Sequence 7240, Ap
15	415	5.1	8189	5 US-09-724-676A-7240	Sequence 7240, Ap
16	415	5.1	9386	5 US-09-724-676-7256	Sequence 7256, Ap
17	415	5.1	9386	5 US-09-724-676A-7256	Sequence 7256, Ap
18	415	5.1	9414	5 US-09-724-676-7239	Sequence 7239, Ap
19	415	5.1	9414	5 US-09-724-676A-7239	Sequence 7239, Ap
20	415	5.1	9466	5 US-09-724-676-7254	Sequence 7254, Ap
21	415	5.1	9466	5 US-09-724-676A-7254	Sequence 7254, Ap
22	415	5.1	9494	5 US-09-724-676-7237	Sequence 7237, Ap
23	415	5.1	9494	5 US-09-724-676A-7237	Sequence 7237, Ap
24	415	5.1	10031	5 US-09-724-676-7255	Sequence 7255, Ap
25	415	5.1	10031	5 US-09-724-676A-7255	Sequence 7255, Ap
26	415	5.1	10059	5 US-09-724-676-7238	Sequence 7238, Ap
27	415	5.1	10059	5 US-09-724-676A-7238	Sequence 7238, Ap
28	415	5.1	10111	5 US-09-724-676-7253	Sequence 7253, Ap
29	415	5.1	10111	5 US-09-724-676A-7253	Sequence 7253, Ap
30	415	5.1	10139	5 US-09-724-676-7236	Sequence 7236, Ap
31	415	5.1	10139	5 US-09-724-676A-7236	Sequence 7236, Ap
32	415	5.1	10356	5 US-09-724-676-7252	Sequence 7252, Ap
33	415	5.1	10356	5 US-09-724-676A-7252	Sequence 7252, Ap
34	415	5.1	10384	5 US-09-724-676-7235	Sequence 7235, Ap
35	415	5.1	10384	5 US-09-724-676A-7235	Sequence 7235, Ap
36	413	5.1	5293	5 US-09-724-676-7107	Sequence 7107, Ap
37	413	5.1	5293	5 US-09-724-676A-7107	Sequence 7107, Ap
38	413	5.1	5626	5 US-09-724-676-7083	Sequence 7083, Ap
39	413	5.1	5626	5 US-09-724-676A-7083	Sequence 7083, Ap
40	412	5.1	5329	5 US-09-724-676-7104	Sequence 7104, Ap
41	412	5.1	5329	5 US-09-724-676A-7104	Sequence 7104, Ap
42	412	5.1	5662	5 US-09-724-676-7054	Sequence 7054, Ap
43	412	5.1	5662	5 US-09-724-676A-7054	Sequence 7054, Ap
44	409	5.1	5329	5 US-09-724-676-7110	Sequence 7110, Ap
45	409	5.1	5329	5 US-09-724-676A-7110	Sequence 7110, Ap

ALIGNMENTS

RESULT 1
US-10-152-319A-2063
; Sequence 2063, Application US/10152319A
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Higgs, Brandon
; APPLICANT: Castle, Arthur
; APPLICANT: Elashoff, Michael
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5089-US
; CURRENT APPLICATION NUMBER: US/10/152,319A
; CURRENT FILING DATE: 2002-05-22
; PRIOR APPLICATION NUMBER: US 60/292,335
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/297,523
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,925
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,810
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/303,807
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/303,808
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/315,047
; PRIOR FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: US 60/324,928
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/330,867
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: US 60/330,462
; PRIOR FILING DATE: 2001-10-22
; Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 2221
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 2063
; LENGTH: 5236
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. NM_057103
US-10-152-319A-2063

Alignment Scores:
Pred. No.: 0 Length: 5236
Score: 8030.00 Matches: 1589
Percent Similarity: 99.75% Conservative: 3
Best Local Similarity: 99.56% Mismatches: 4
Query Match: 99.47% Indels: 0
Gaps: 0
DB:
US-09-902-432-4 (1-1596) x US-10-152-319A-2063 (1-5236)

QY 1 MetGlyAlaGlySerSerThrGluGlnArgSerProGluGlnProAlaGlySerAspThr 20
DB 33 ATGGGGCGCAGGCGAGTTCACCGGACGCGAGGCGCGAGCGCGGCGGAGCGACAG 92
QY 21 ProSerGluLeuValLeuSerGlyHisGlyProAlaAlaGluAlaSerGlyAlaAlaGly 40
DB 93 CCGAGGAGACGTGGTCTCAGTGGCCATGGGCGCGCAGCTGAAAGCTCGGGAGCAGTGA 152
QY 41 AspProAlaAspAlaAspProAlaThrIleuProGlnIleuAsnGlyIleuSerSer 60
DB 153 GACCCCGCGCAGCGGAGCCCGCCGACCAAGCTCCACAGAAAGATGGCCAGCTGCTTCT 212
QY 61 ValAsnGlyValAlaGluGlnGlyAspValHisValGlnGluGluAsnGlnGlyGln 80
DB 213 GTCAACGGCGGTGATGAAACAGAGAGATGTCCATGTCCAAAGAGGAAACCAAGAGGGGCGAG 272
QY 81 GluGluGluValValAspGluAspValGlyGlnArgGluSerGluAspValArgGlyIle 100
DB 273 GAGGAAGAAAGCTGTGATAGATGTGGACAGCGAGCTCAGAAAGATGTGAGAGAAAAA 332
QY 101 AspArgValAlaGluIleuMetAlaAlaAsnSerThrAlaValAlaGluAspIleThrIleuAspGly 120
DB 333 GACCGAGTTGAAGAAATGGCGGCGCACTCCACGCTGTGAAGATATCCAAAGATGGG 392
QY 121 GlnGluGlnThrSerGluIleIleGluGlnIleProAlaSerGluAsnAsnValGluGlu 140
DB 393 CAGGAGGAGACATCAGAAATTAATTGAACAGATCCCTGCTTCAGAAAAACATGGAGAA 452
QY 141 MetValGlnProAlaGluSerGlnAlaAsnAspValGlyPheIleuValPheIleuPhe 160
DB 453 ATGTGACAGCTGTGAGTCCAGGCTAAATGATGTGGCTTCAAGAAAGTATTTAATT 512
QY 161 ValGlyPheIleuPheThrValIleuIleuAspIleuAsnGluIleuSerSerThrValGlnLeu 180
DB 513 GTTGTGTTTAAATTCACGCGTGAAGAAAGATTAATAAGTAAGTCACTGTCCAACTA 572
QY 181 LeuThrValIleuValAspGluGlyGluGlyAlaGluAlaSerValGlyAlaGlyAspHis 200
DB 573 CTCACGTGTCAAGAGGATGAAGCGCAAGAGGCGCAAGAGCTCTGTCCGAGCTGAGAC 632
QY 201 GlnGluProSerValGluThrAlaValAlaGlyIleuSerAlaSerIleuSerGluIleuLeu 220
DB 633 CAGGAGCCCAAGTGTGAGACTGCTCCGAGAGTCCAGATCCAAAGAAAGTGAAGTGAAG 692
QY 221 GlnSerThrGluIleuGlnGluGlyThrIleuValGlnGlnIleuSerSerThrGluIlePro 240
DB 693 CAATCCACAGAGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 752
QY 241 LeuGlnAlaGluSerAspGlnAlaAlaGluGluGluValAlaIleuAspGluGlyGluGly 260
DB 753 CTTCAAGCGCAATCTGATCAAGGGGCTGAGGAAGAGCCAAAGATCAAGAGAGAGAAAA 812
QY 261 GlnGluIleuValValSerThrIleuSerProGluSerProSerSerProValIleuSerGluThr 280

DB 813 CAGAGAAAGACCCCAAGTCCCAAGATCCCGAGACCCCAAGTCAACAGTATAGCA 872
QY 281 ThrSerSerPheIleuValPhePheThrHisGlyTyrAlaGlyTyrPargIleuValSer 300
DB 873 ACATCTTCTTCAAGAAATTTCTTCACTCAGCGTTGGCCGGCTGGCGCAAGAGCACCG 932
QY 301 PheIleuValSerValGluAspAlaGluGlnThrAlaGluValSerGlyGluGlnGluVal 320
DB 933 TTCAGAAATCAAAAGAGATGATCTGAAACTGCGGAGAGAGAAAGAGCAAGCAAGCA 992
QY 321 GluIleuValAspGluGluGluIleuSerGluIleuSerThrGluProAlaSerGluGlnGluPro 340
DB 993 GAAAAAGTACAG 1052
QY 341 AlaGluAspThrAspGlnAlaArgIleuSerAlaAspTyrGluIleuValGluIleuPro 360
DB 1053 GCAGAAAGACACAGACCAAGGCGGTTTCAAGACACTACAGAAAGTGGAGCTGCTTTG 1112
QY 361 GluAspGlnValGlyAspLeuGluAlaSerSerGluGluValSerValProLeuAlaThr 380
DB 1113 GAAAGACAGGTTGTGACCTGGAGCATGTCAAGAGAAAGTGTGCTTTGGCAACG 1172
QY 381 GluValPheAspGluIleuMetGluAlaHisGlnGluValAlaGluValHisValSer 400
DB 1173 GAACTTTTGATGAGAAATGAGAACCCACCAAGAGTTGTGGCAGAGGTCACGTTGAGC 1232
QY 401 ThrValGluIleuSerThrGluGluGlnGlyGlyGlyGlyGluAlaGluGlyValVal 420
DB 1233 ACCGTGAGAAAGACAG 1292
QY 421 ValGluGlyThrGlyIleuSerLeuProProGluIleuValGluProGlnGluValPro 440
DB 1293 GTAGAGAGAAACAGAGAAATCTTGGCCCTGAGAAACTGGCTGAGAGAGAGAGAG 1352
QY 441 GlnGluAlaGluProAlaGluGluIleuMetIleuSerSerArgGluMetCysValSerGly 460
DB 1353 CAGAAAGCTGAGCTGTGAGAGGCTGATGAAGACCAAGAGATGTGTCTGTGAGAGA 1412
QY 461 AspHisThrGlnLeuThrAspLeuSerProGluGluIleuValSerThrLeuProIleuHisProGlu 480
DB 1413 GACCACTCAACTCAGACAGACCTTAAGTCTTGAAGAGAGAGAGAGAGAGAGAGAG 1472
QY 481 GlyIleuValSerGluValGluMetLeuSerSerGlnGluGlyIleuValGlnGlySer 500
DB 1473 GGCATTGTCAAGAGAGTGAAGATGTCTGTCTCAGAAAGATCAAGTCAAGAGAA 1532
QY 501 ProLeuIleuValLeuPheSerSerSerGlyLeuIleuValSerGlyIleuValSerGly 520
DB 1533 CCTTGAAAGAACTCTTCAAGTACCTCAGGCTTAAAGAAAGCTGTCTGGAGAGAGAG 1592
QY 521 GlyIleuValSerGlyGlyGlyAspGluGluProGluGlyIleuValSerGlyIleuValSer 540
DB 1593 GGGAAACGAGAGAGGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1652
QY 541 SerProGluSerAlaAspGluGlnIleuValGlyIleuSerAlaSerSerProGluIleuPro 560
DB 1653 TCCCAAGAGAGTGTGAGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1712
QY 561 GluGluThrThrCysLeuGlnIleuGlyProLeuGluAlaProGlnIleuValGluAla 580
DB 1713 GAGGAGACCAAGTGTCTGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1772
QY 581 GluGlyThrThrSerAspGlyGlyIleuValSerGlyIleuValSerProTyrAlaSerPhe 600
DB 1773 GAAGGAATCTACTTCCGAG 1832
QY 601 LysIleuMetValThrProIleuValSerGlyValArgArgProSerGluSerAspGlyGlu 620
DB 1833 AAAAAGATGTGACACCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1892
QY 621 GluLeuGluValValSerSerAlaThrIleuSerThrAspSerThrValSerGluMet 640

Db 1893 GAGCTGGAGAGGTCAGAGCGCCACCTTGTCTCCTCCTAGTACGACAGTGTGAGAAATG 1952
Qy 641 GluAspGluValIysThrValGlyGluGluGlnIysProGluGluProIysArgArgVal 660
Db 1953 CAAGATGAAGTCAAAACTGTTGGTGAGGAAACAAAGCCAGAGAAACCAAAGCGTAGGGTG 2012
Qy 661 AspThrSerValSerTrpGluAlaLeuIleCysValGlySerSerIysIysArgAlaArg 680
Db 2013 GATATCTCAGTGTCTTGGGAAGCACTGATTGTGTGGATCATCCCAAGAGAGCAAGG 2072
Qy 681 LysAlaSerSerSerAspAspGluGlyGlyProArgThrLeuGlyGlyAspSerHisArg 700
Db 2073 AAGCATCTCTTCAGATGATGAAGAGGCCAAGGACACTGGGAGGGACAGTCACAGA 2132
Qy 701 AlaGluAlaSerIysAspIysGluAlaGlyThrAspAlaValProAlaSerThrGln 720
Db 2133 GCAGAGAGGCCAGCAAGAACAAAGAACCGGAACAGACGCTGTCTCTGCAGCACCCAG 2192
Qy 721 GluGlnAspGlnAlaGlnGlySerSerSerProGluProAlaGlySerProSerGluGly 740
Db 2193 GAGCAGACCCAGCGCAGGAAGGTTCTCACCAGGACCGGGAGGCCCTTCGGAAGGG 2252
Qy 741 GluGlyValSerThrTrpGluSerPheLysArgLeuValThrProArgIysLysSerIys 760
Db 2253 GAAGGTGTCTCCACTTGGGAGTCATTTAAAGATTAGTCACTCCAGAAAAAATCCAG 2312
Qy 761 SerLysLeuGluGluIysAlaGluAspSerSerValGluGlnLeuSerThrGluIleGlu 780
Db 2313 TCAAACTGGAAGAGAAAGCCGAAGACTCTAGTGTAGACAGTTGTCCACTGAGATCGAA 2372
Qy 781 ProSerArgGluGluSerTrpValSerIleIysIysPheIleProGlyArgArgLysLys 800
Db 2373 CCAGGTAGAGAAGAAATCTTGGGTTTCCATTAAAGAAATTCATCCCGGACCGCGGAAGAA 2432
Qy 801 ArgAlaAspGlyIysGlnGluGlnAlaThrValGluAspSerGlyProValGluIleAsn 820
Db 2433 AGGCAGATGGGAAGCAAGAACCAAGCCACTGTGTGGAAGACTCAGGCCACGTGGAGTAAT 2492
Qy 821 GluAspAspProAsnValProAlaValProLeuSerGluTrpAsnAlaValGluArg 840
Db 2493 GAGGACGACCTGATGCCAGCGCTGCTGCTCTGCTGATGATGATGATGATGGAGG 2552
Qy 841 GluIysMetGluAlaGlnGlyAsnThrGluLeuProGlnLeuLeuGlyAlaValTyrVal 860
Db 2553 GAGAAGATGAAGCCAGGGGAAATCGCGAGCTGCCCCAGCTGTGGGGGCTGTGTACGTG 2612
Qy 861 SerGluGluLeuSerLysThrLeuValHisThrValSerValAlaValIleAspGlyThr 880
Db 2613 TCCAGAGGCTCAGTAAGACTCTGGTCCACACTGTGAGTGTCCGAGTCATTTGATGGACC 2672
Qy 881 ArgAlaValThrSerValGluGluArgSerProSerTrpIleSerAlaSerValThrGlu 900
Db 2673 AGGCAGTCAACAGTGTGGAAGAGCGGTCTCTTCTGTGGATATCCGCTTCGGTAACAGAA 2732
Qy 901 ProLeuGluHisThrAlaGlyGluAlaMetProProValGluGluValThrGluLysAsp 920
Db 2733 CCTCTTTGAACACACAGCGGAGAACCCATGCCACTGTGAAGAGGTCACTGAAAAAGAC 2792
Qy 921 IleIleAlaGluThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHis 940
Db 2793 ATCATTCAGAAAGAACTCTCTGTCTCACCAGACGTTACAGAGGGTAAGATGCCCAT 2852
Qy 941 AspAspMetValThrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThr 960
Db 2853 GACGACATGTTCACCATGAAGTGAATTTACCTCAGAACTGTGACAGGCCACAGAGACC 2912
Qy 961 SerGluAlaLeuArgThrGluGluValThrGluAlaSerGlyAlaGluGluThrThrAsp 980
Db 2913 TCAGAGGCTCTCCGTACTGAAGAAGTTACCGAAGCATCGGGGGCCGAAGAGACCACAGAC 2972
Qy 981 MetValSerAlaValSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrPro 1000
Db 2973 ATGTTGTCGCGAGTTTCCCAAGCTGACTGCTCCCAAGACTTCCCAAGAGGAAAGCCCA 3032

Qy 1001 ValGlnGluValGluSerGlyValLeuAspThrGluGluGluArgGlnThrGlnAla 1020
Db 3033 GTTCAGAGGTAGAGAGTGGTGTGTAGATACAGAAAGAGAGAGCGCCAGAGCGAGCC 3092
Qy 1021 IleLeuGlnAlaValAlaAspIysValIysGluGluSerGlnValProAlaThrGlnThr 1040
Db 3093 ATCCTCCAAGCCGTTGAGACAAGGTGAAGAGAGTCCAGGTGCTTGCACCCAGACT 3152
Qy 1041 ValGlnArgThrGlySerLysAlaLeuGluLysValGluValGluGluAspSerGlu 1060
Db 3153 GTCAGAGAACCGGGTCNAAGACACTGGAGAGGTTGAGGAGGTAGAGAGGACTCCGAA 3212
Qy 1061 ValLeuAlaSerGlnIysGluLysAspValMetProIysGlyProValGlnGluAlaGly 1080
Db 3213 GTGCTGGCTTCGAGAGAAAGAGAGGACGTTATGCCAAAAGAGCCCGTGCAAGAGCTGA 3272
Qy 1081 AlaGluHisLeuAlaGlnGlySerGluThrGlyGlnAlaThrProGluSerLeuGluVal 1100
Db 3273 GCTGAGCATCTTCACAGGGCTCTGAGACTGGACAGGCTACTCCAGAGAGCCTTTGAAGTT 3332
Qy 1101 ProGluValThrAlaAspValAspHisValAlaThrCysGlnValIleLysLeuGlnGln 1120
Db 3333 CCTGAAGTCACGCGCAGATGTAGACCATGTGCCACGTCGAGGTTATCAAGCTCCAGCAG 3392
Qy 1121 LeuMetGluGlnAlaValAlaProGluSerSerGluThrLeuThrAspSerGluThrAsn 1140
Db 3393 CTGATGGAAACAGCGCGCTGGCCCTGAGTCATCCGAAACCTTGACAGACAGTGAGACAAAT 3452
Qy 1141 GlySerThrProLeuAlaAspSerAspThrAlaAspGlyThrGlnGlnAspGluThrIle 1160
Db 3453 GGAAGCACTCCCTTACAGATTCCAGACTGCAGATGGGACACAGCAAGATGAACCATTT 3512
Qy 1161 AspSerGlnAspSerLysAlaThrAlaAlaValArgGlnSerGlnValThrGluGluGlu 1180
Db 3513 GACAGCCAGGACAGTAAGCCACTGCAGCTGTCCAGCAGTCACAGGTCACAGAAAGAGAG 3572
Qy 1181 AlaAlaThrAlaGlnIysGluGluProSerThrLeuProAsnAsnValProAlaGlnGlu 1200
Db 3573 CGCGCTACTGCTCAGAAAGAGGAGCCTTCGACACTACTTAATAATGTTCCAGGCCAGAA 3632
Qy 1201 GluHisGlyGluGluProGlyArgAspValLeuGluProThrGlnGlnGluLeuThrAla 1220
Db 3633 GAACATGGGGAGAACAGCAGAGAGAGATGTTCTTGAACCTTACACAGCAAGAGCTTACTGCT 3692
Qy 1221 AlaAlaValProValLeuAlaLysThrGluValGlyGlnGluGlyGluValAspTrpLeu 1240
Db 3693 CGAGCGGTGCCCTTCTGGCAAGACTGAGTGGGTCAAGAGGGTGAGGTTGACTGGTTG 3752
Qy 1241 AspGlyGluLysValIysGluGlnGluValPheValHisSerGlyProAsnSerGln 1260
Db 3753 GATGGAGAAAAAGTCAAAAGAAACAGGAGGTGTTGTACACTCTGGACCCCAACAGTCAA 3812
Qy 1261 LysAlaAlaAspValThrTyrAspSerGluValMetGlyValAlaGlyCysGlnGluLys 1280
Db 3813 AAGGCTGCTGATGTGACATATGACAGTGAAGTCAAGGAGTGGCCGGGTGTCTGAGAAAG 3872
Qy 1281 GluSerThrGluValGlnSerLeuSerLeuGluGluGluMetGluThrAspValGlu 1300
Db 3873 GAGGTACTGAAGTGCAGAGTCTTAGCTGGAGAGGGAGAGATGGAATCTGACGTTGAA 3932
Qy 1301 LysGluLysArgGluThrLysProGluGlnValSerGluGlyGluGlnGluThrAla 1320
Db 3933 AAGGAGAAAGGGAGACAAAGCCAGAGCAAGTCAAGTGAAGAGGTGAGCAGGAAACAGACC 3992
Qy 1321 AlaProGluHisGluGlyThrTyrGlyLysProValLeuThrLeuAspMetProSerSer 1340
Db 3993 GCTCCTGAGCATCAGGAACCTACGGGAAGCCAGTCTGACACTTGACATTCGCCAGCTCA 4052
Qy 1341 GluArgGlyLysAlaLeuGlySerLeuGlySerProSerLeuProAspGlnAspLys 1360
Db 4053 GAGAGGGGGAAGGCACTGGGAAGCCTTGGAGGAAGCCCTTCTCTCCAGAGCAAGCAAA 4112

```

Oy 1361 ALaGlyCysIleGluValGlnValGlnSerLeuAspThrThrValThrGlnThrAlaGln 1380
Db 4113 GAGGTTCATAGAGATTCATAGTTCATAGCTGAGACACAGCTCATCAACACGACGAA 4172
Oy 1381 ALaValGluValValIleGluThrValValIleSerGluThrGlyGluSerProGluCys 1400
Db 4173 GCTGGGAAAAGTCATAGAAACGGTTGATTTCAGACACAGCGTGAAGTCCAGAGTGT 4232
Oy 1401 ValGlyAlaHisLeuLeuProAlaGluLysSerSerAlaThrGlyGlyHisThrPheLeu 1420
Db 4233 GTAGGTGACACTTATTCACACCTGAGAGTCTCTGCAACCGGTGGCCACCTGACTCTT 4292
Oy 1421 GlnHisAlaGluAspThrValProLeuGlyProGluSerGlnAlaGlnSerIleProIle 1440
Db 4293 CAGCATGACAGAGACACGGTACCCCTGGGGCTGAGTCTCAGGACGAATCCATCCCAATC 4352
Oy 1441 IleValThrProAlaProGluSerThrLeuHisProAspLeuGlnGlyGluIleSerAla 1460
Db 4353 ATAGTACTCTGCTCTCTTAAAGCACCCCTACATCTGACCTTCAGAGGAATTAAGCGCA 4412
Oy 1461 SerGlnArgGluArgSerGluGluGluAspLysProAspAlaGlyProAspAlaAspGly 1480
Db 4413 TCCACAGAGAGACCGATCCAGAGAGAGCAAGCCAGATGCTGCTCTGATCTGACGCGC 4472
Oy 1481 LysGluSerThrAlaIleGluLysValLeuLysAlaGluProGluIleLeuGluLeuGlu 1500
Db 4473 AAGGAGAGTACAGCATCAAAAAGCTCTCAAGGCTGAACCTGAGTCTTGAACTTGAG 4532
Oy 1501 SerLysSerAsnLysIleValLeuAsnValIleGlnThrAlaValAspGlnPheAlaArg 1520
Db 4533 AGTAAGACGAACAAGATTGCTGTAACGTCATTCAGACACCGCTTGACCAAGTTCGACGT 4592
Oy 1521 ThrGluThrAlaProGluThrHisAlaTyrAspSerGlnThrGlnValProAlaCysArg 1540
Db 4593 ACAGAAACAGCCCGCAAACTCATGCTTATGATTCACAGACCCAGGTTCCTGCAATGGCAG 4652
Oy 1541 LeuAspSerArgGluProAsnArgCysTyrPheLysMetLysAspAlaLysMetLysHis 1560
Db 4653 GCTGACAGAGGAGCCCAACAGATCTGACAAAAATTAAGTTGCCAAGATGAACAC 4712
Oy 1561 ProValProGlnProArgGluAspLeuGlnValLeuThrValLeuGlnValLeuAlaGln 1580
Db 4713 CCAGTCCCGCCACCCAGAGAGACTTGGCAAGTCTGACCGTCTCTGAGGACATGGGCTCAG 4772
Oy 1581 ProArgLysCysLeuProArgLeuGlnLeuLysAlaProValSerLys 1596
Db 4773 CCTCGGAATGCTTGCGCGCTTGCAATTGAAAGCGCGGTGTCAAG 4820

RESULT 2
US-09-724-676-12735
; Sequence 12735, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; NUMBER OF SEQ ID NOS: 2009-11-28
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12735
; LENGTH: 5931
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676-12735

Alignment Scores:
Pred. No.: 3,65e-172 Length: 5931
Score: 4341.50 Matches: 970
Percent Similarity: 69.13% Conservative: 190
Best Local Similarity: 57.81% Mismatches: 110
Query Match: 53.78% Gaps: 33

```

```

US-09-902-432-4 (1-1596) x US-09-724-676-12735 (1-5931)
Oy 1 MetGlyAlaGlySerSerThrGluGlnArgSerProGluGln---ProAlaGlySerAsp 19
Db 192 ATGGCGCGCGGAGAGCTCCACCGAGAGCGCAGCCCGGAGCAGCCCGAGGGGAGCTCC 251
Oy 20 ThrProSerGluLeuValLeuSerGlyHisGlyProAlaAlaGluAlaSer---GlyAla 38
Db 252 AGCGCGGCTGAGCCCGAGCCCGACCCAGCGCGCGCGCTCTGGCGGAGGCGCGCACACCC 311
Oy 39 AlaGlyAspProAlaAspAla-----AspProAlaThrLysLeuProGlnLysAsnGly 56
Db 312 ACCCGGAGCCCGCATGCTGCTCGGACCCCGCACCAAGCTCTTCAGAGAATGT 371
Oy 57 GlnLeuSerSerValAsnGlyValAlaGluGlnGlyAspValHisValGlnGluGln 76
Db 372 CAGCTGTCCACATCAATGAGCGTAGCTGAGCAAGATGAGTCCAGCTCCAGAGGCTGAC 431
Oy 77 GlnGluGlnGln-----GlnGlu 82
Db 432 CTAAATGGCCAGAAAGAGCCCTGAACTGAGAGGACCTTAAACAGCCAGAGAGAAAGAA 491
Oy 83 GluValValAspGluAspValAlaGlyGlnArgGluSerGluAspValArgLysAspArg 102
Db 492 GAAGTCAATTGCACAGAGCTTGAGACAGAGACTTGAAAGATGTGAGCGGAAGAGACTCC 551
Oy 103 ValGluGluMetAlaAlaAsnSerThrAlaValGluAspIleThrLysAspGlyGlnGlu 122
Db 552 GATAAAGAGATGGCTTACTAAGTACCGGTTGTTCCAGCATCACAGATGATGGCAGAG 611
Oy 123 GlnThrSerGluIleIleGlnGlnIleProAlaSerGluAsnAsnValGlnGluMetVal 142
Db 612 GAGACACCCGAAATATATGAACAGATTCCTCTTCAAGAACCAATTATGAAGAGCTAA 671
Oy 143 GlnProAlaGluSerGlnAlaAsnAspValAlaPheLysLysValPheLysPheValGly 162
Db 672 CAACCCACGTAGTCCACAGCTTATGATGATTTAAGAAAGTGTTTAAGTTCTTGCG 731
Oy 163 PheLysPheThrValLysValAspLysAsnGluLysSerAspThrValGlnLeuLeuThr 182
Db 732 TTTAAATTCACTGTGAAAAAGATTAAGACAGAGAACTGACACTGTCCAGTACTACT 791
Oy 183 ValLysLysAspGluGluGluGluValAlaGluAlaSerValGlyAlaGlyAspHisGlnGlu 202
Db 792 GTAAAGAAAGATGAAGGGGAGGAGCA-----GAGGGGCTGGGACCAACAGGAG 842
Oy 203 ProSerValGluThrAlaValAlaGlyLysSerAlaSerLysGluSerGluLeuLysGlnSer 222
Db 843 CCCAGCTT-----GGGGCTGGAGAGACAGCTCCAAAGAAAGGAAACCCAAACATCT 896
Oy 223 ThrGluLysGlnGluGlyThrLeuLysGlnGlnGlnSerSerThrGluIleProLeuGln 242
Db 897 ACAGAGAAACCCGAGAGACCTGTAAGCGTAGCAACCAACGCAAGATTTCTCCCCCA 956
Oy 243 AlaGluSerAspGlnAlaAlaGluGlnGluLysAspGluGluGluLysGlnGlnGln 262
Db 957 GCGGAATCTGGCCAGACAGTG--GAGGAATGCCAAAGGAGGAAGGAAGAAACAAAGAA 1013
Oy 263 LysGluProThrLysSerProGluSerProSerSerProValAsnSerGluThrThrSer 282
Db 1014 AAAGAACTAGCAAGTCTGCAGAAATCTCCGACTAGTCCCGTGCACAGTGAAGAGAGATCA 1073
Oy 283 SerPheLysLysPhePheThrHisGlyTyrAlaGlyTyrPargLysLysThrSerPheLys 302
Db 1074 ACCTTCAAAAATTTCTTCACTCAAGGTTGGCGCGCTGCGCAAAAAGACCAAGTTTCAGG 1133
Oy 303 LysSerLysGluAspAspLeuGluThrAlaGluLysArgLysGluGlnGlnAlaGluLys 322
Db 1134 AAGCCGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1193
Oy 323 ValAspGluGluGluLysGluLysThrGluProAlaSerGluGlu----- 337
Db 1194 GTAGACACAGAAAGAAAGCGAAAGGACAGAGGTTGCTCCGAGAAACTGACCGCTCCGAG 1253

```

QY 338 -----GlnGluProAlaGluAspThrAspGlnAlaArgLeuSerAlaAspTyr 353
Db 1254 CAAGCCACCACAGGAGCGCGAGAAAGTGCACAGCGCCGGTTATCAGCTGAATAT 1313
QY 354 GluLysValGluLeuProLeuGluAspGlnValGlyAspLeuGluAlaSerSerGluGlu 373
Db 1314 GAGAAAGTGGAGTGCCTCAGAGGAGCAAGTCAGTGGCTCGCGGACCTTCTGAAGAG 1373
QY 374 LysCysAlaProLeuAlaThrGluValPheAspGluLysMetGluAlaHisGln---Glu 392
Db 1374 AAACCTGCTCGTGGCGCAGAGAAGTGTGATGAGAAAATAGAAGTCCACCAAGAGAG 1433
QY 393 ValValAlaGluValHisValSerThrValGluLysThrGluGluGlnGlnGlyGly 412
Db 1434 GTTGTGGCCGAAGTCCAGCTCAGCACCGTGGAGAGAGAACCCGAAGAGCAG----- 1484
QY 413 GlyGluAlaGluGlyValValValGluGlyThrGlyGluSerLeuProGluLys 432
Db 1485 -----AAAACGGAGGTGGAAGAAACACAGCGGTCTGTGCCAGCTGAAGAA 1529
QY 433 LeuAlaGluProGlnGluValProGlnGluAlaGluProAlaGluLeuMetLysSer 452
Db 1530 TTGTTGAAATGGATGCAGAACTCAGAAAGCCGAACCTGCCAAGGAGCTGGTCAAGCTC 1589
QY 453 ArgGluMetCysValSerGlyGlyAspHisThrGlnLeuThrAspLeuSerProGluGlu 472
Db 1590 AAAGAAACGTGTGTTCCGAGAGGAGACCTACACAGGAGCTACCTCAGTCTGTATGAG 1649
QY 473 LysThrLeuProLysHisProGluGlyLeValSerGluValGluMetLeuSerSerGln 492
Db 1650 AAGTGTCTGCCAAACCCCGAAGGGGTGTGTAGTGAGGTGGAATGTCTCATCACAG 1709
QY 493 GluArgIleLysValGlnGlySerProLeuLysLysLeuPheSerSerSerGlyLeuLys 512
Db 1710 GAGAAATGAAGTGCAGGAGAGTCCACTAAGAGAGCTTTTACCAGACTGGCTTAAAA 1769
QY 513 LysLeuSerGlyLysLysGlnLysGlyLysArgGlyGlyGlyAspGluGluProGly 532
Db 1770 AAGCTTTCTGAAAGAAACAGAAAGGGAAGA---GGAGGAGGAGACGAGGAATCAGGG 1826
QY 533 GluTyrGlnHisIleHisThrGluSerProGluSerAlaAspGluGlnLysGlyGluSer 552
Db 1827 GAGCACACTCAGGTTCAGCGCATTTCCCGACAGCAGGAGGAGGAGAAAGGGCGAGAGC 1886
QY 553 SerAlaSerSerProGluGluProGluGluThrThrCysLeuGluLysGlyProLeuGlu 572
Db 1887 TCTGCCTCATCCCTGAGGAGCCGAGAGATCAGCTGTCTGGAAAAGGGCTTAGCCGAG 1946
QY 573 AlaProGlnAspGlyGluAlaGluGluGlyThrThrSerAspGlyGluLysLysArgGlu 592
Db 1947 GTGCAGCAGGATGGGAAGCTGAAGAAGGAGTACTTCCGATGGAGAGAAAGAAAGAA 2006
QY 593 GlyIleThrProTrpAlaSerPheLysMetValThrProLysLysArgValArgArg 612
Db 2007 GGTGTACTCTCCCTGGGCATCATTTCAAAGAATGTGTACGCCCAAGAACGCTGTAGACG 2066
QY 613 ProSerGluSerAspLysGluGluGluLeuLysValLysSerAlaThrLeuSerSer 632
Db 2067 CCTTCGGAAGTGAATAAGAGATGAGCTGGACAGGTCAAGAGCGCTACTCTGTCTTCC 2126
QY 633 ThrAspSerThrValSerGluMetGlnAspGluValLysThrValGlyGluGlnGlnLys 652
Db 2127 ACCGAGAGCAGCCTCTGAAATGCAAGAAATGAAGGAGCGTGGGAAGGCCCAAAG 2186
QY 653 ProGluGluProLysArgArgValAspThrSerValSerTrpGluAlaLeuIleCysVal 672
Db 2187 CCGAAGAACCAAGCCCAAGGTGGATACCTCAGTATCTTGGGAAGCTTTAAATTTGTGTG 2246
QY 673 GlySerSerLysLysArgAlaArgLysAlaSerSerSerAspGluGlyGlyProArg 692
Db 2247 GGATCATCCAAAGAAAGCAAGGAGGAGGTCTCTTCTGATGAGGAAGGGGAGCAAA 2306

QY 593 ThrLeuGlyGlyAspSerHisArgAlaGluAlaSerLysAspLysGluAlaGlyThr 712
Db 2307 GCAATGGGAGGAGACACCCAGAAAGTGTATGAGCGCGGAAACAAAGACAGCGGACA 2366
QY 713 AspAlaValProAlaSerThrGlnGluGlnAspGlnAlaGlnGlySerSerProGlu 732
Db 2367 GACGGATCTTCTGCTGGTCCCAAGAACATGATCCAGGCGCAGGAAGTTCTCCCCGGAG 2426
QY 733 ProAlaGlySerProSerGluGlyGlyValSerThrTrpGluSerPheLysArgLeu 752
Db 2427 CAAGCTGGAAGCCCTACCAAGAGGGAGGGCTTTCCACCTGGGAGTCATTTAAAGAGTTA 2486
QY 753 ValThrProArgLysSerLysSerLysLeuGluGluLysAlaGluAsp----- 769
Db 2487 GTCACGCCCAAGAAAAAATCAAAGTCCAAGCTGGAAGAGAAAAAGAAAGACTCCATAGCT 2546
QY 770 ---SerSerValGluGlnLeuSerThrGluIleGluProSerArgGluGluSerTrpVal 788
Db 2547 GGTCTGTGTAGAACATTTCCACTCCAGACACTGAACCCCGTTAAAGAGNANTCTGGGTC 2606
QY 789 SerIleLysPheIleProGlyArgArgLysArgAlaAspGlyLysGlnGln 808
Db 2607 TCAATCAAGAAAGTTTATCTCGACCAAGGAAGAAAGGCCAGATGGGAAACAAGAACAA 2666
QY 809 AlaThrValGluAspSerGlyProValGluIleAsnGluAspAspProAsnValProAla 828
Db 2667 GCCCTCTGTGAAGACACGCGGCCAACAGGGGCCCAACGAAGATGACTCTGTATGTCCCGGC 2726
QY 829 ValValProLeuSerGluTyrAsnAlaValGluArgGluLysMetGlu-----Ala 845
Db 2727 GTGTCCTCTGTCTGAGTATGATCTGTAGAAAGGAGAAATGGAGGCACAGCAAGGCC 2786
QY 846 GlnGlyAsnThrGluLeuProGlnLeuLeuGlyAlaValTyrValSerGluGluLeuSer 865
Db 2787 CAABAAGCCGACAGCAGCCGAGCAGAGGAGCCACTGAGGTGTCAGAGGAGCTCAGC 2846
QY 866 LysThrLeuValHisThrValSerValAlaValIleAspGlyThrArgAlaValThrSer 885
Db 2847 GAGAGTCAGGTTCATATGATGGCAGCAGCTGTCTGACGGGACGAGGCGAGCTTACATT 2906
QY 886 ValGluGluArgSerProSerTrpIleSerAlaSerValThrGluProLeuGluHisThr 905
Db 2907 ATTGAAGAAGGTCTCTCTTGTGATATCTGTTCAAGTACAGAACCTCTTGAACAAGTA 2966
QY 906 AlaGlyGluAlaMetProProValGluGluValThrGluLysAspIleLeAla---Glu 924
Db 2967 GAAGCTGAAGCCGACACTGTTAACTGAGGAGGTATTGGAAGAGAGTAATTCAGAGAAGAA 3026
QY 925 GluThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHisAspAspMetVal 944
Db 3027 GAACCCCCACCGTTACTGAACCTCTGCCAGAGAACAGAGAGCGCCGGCGCACCGCTC 3086
QY 945 ThrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeu 964
Db 3087 GTTAGTGAGCGGAATTTGACCCCGAAGCTGTGACAGCTGCAGAAACTGCAGGCCCATTTG 3146
QY 965 ArgThrGluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSerAla 984
Db 3147 GGTGCCGAGAGAGAACCCGAGGATCTGTCTGTGAAGAGACCAAGAAATGGTGTACGCA 3206
QY 985 ValSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGluVal 1004
Db 3207 GTCTCCAGTTAACCGACTCCCCAGACACCCAGAGAGGAGGCCACTCCGGTGCAGGAGGTG 3266
QY 1005 GluSerGlyValLeuAspThrGluGluGluArgGlnThrGlnAlaIleLeuGlnAla 1024
Db 3267 GAAAGTGGCGTACTCAGCATAGAGAGCAAGAGAGCGGACTCAAGAGGTCTCTCCAGGCA 3326
QY 1025 ValAlaAspLysValLysGluGluSerGlnValProAlaThr----- 1038
Db 3327 GTGGCAGAAAGTGAAGAGGAAATCCAGCTCTGGCACCGCTGGGCCAGAGATGTG 3386
QY 1039 ---GlnThrValGlnArgThrGlySerLysAlaLeuGluLysValGluGluValGluGlu 1057


```

Db 3387 CTTCAGCTGTGCAGAGACA-----GAGCGAGAAAGACCAAGAG 3428
Qy 1058 AsperglutalLeuAlaserGluysgluysAspValMetProlysglyProvalGln 1077
Db 3429 CAGCGTGAAGCGTGGGCTGTGAAGAAAGACGAGATGTGTGTAAGTAGATGCTCAG 3488
Qy 1078 GluAlaGlyAlaGluHisLeuAlaserGluThrGlyGlnAlaThrProGluSer 1097
Db 3489 GAGCGAAAAACTGACCTTTTACACAGAGAGAGGTGTGGGACAGCCAGGAAAC 3548
Qy 1098 LeuGlu-----ValProGluValThrAlaAspValAspHis-----ValAlaThrCys 1113
Db 3549 TTTGAAAAAGCTCTCTCAATGTCACAGAGACATAGACTCCAGTGAAGCTTTAACCCTTGT 3608
Qy 1114 GluVal-----LysLeuGlnGlnLeu-----MetGluGlnAlaValAla 1127
Db 3609 CAGCGCAAACTTACGTGGGTAAATACAGAGAGATGTGTGAAACAGGCTATATCCCC 3668
Qy 1128 ProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrProLeuAlaAsp 1147
Db 3689 CCGTACTCGGTGAAACCCCTTACAGACGTGAGACTGATGAGAACCCCGGTAGCCGAC 3728
Qy 1148 SerAspThrAlaAspGlyThrGlnGlnAspGluThrAlaAspSerGluAspSerLysAla 1167
Db 3729 TTTGACGACCCAGGACCAACCAAGAAAGCAAGATTGTGAATCCATAGAGAAATGAG 3788
Qy 1168 ThrAlaAlaValArgLysSerGlnValThrGluGluGluAlaAlaThrAlaGlnGlu 1187
Db 3789 GTCCGATCTCGTACCAAGTCAAGGGGCAAGAGACAGAGAGCTTCTGCAAGAAAGAG 3848
Qy 1188 GluProSerThrLeuProAsnAsnValProAlaGlnGlnGluHisGlyGluGluProGly 1207
Db 3849 AGGCTCCAGCA-----CTTCCAGTTTGTGTTCCAGGAAACCTAAGAACATCCAAAG 3905
Qy 1208 ---ArgAspValLeuGluProThrGlnGlnGluLeuThrAlaAlaAlaValProValLeu 1226
Db 3906 ATGAGAGACCTCTAGACATACAGATTAAGAGGTCTAGTGAACCTATATTCATTCTG 3965
Qy 1227 AlaLysThrGluValGlyGlnGluGluValAspTrpLeuAspGlyGluLysValLys 1246
Db 3966 TCAAGACGAGAGGAGCTCA-----GAGGCTGACCAATATGCTGATGAGAAACAA 4019
Qy 1247 -----GluGlnGlnGluValPheValHisSerGly-----ProAsnSer 1259
Db 4020 GACGTACCATTTTTCAGAGACTTGGAGGCTGTATGACACAGGCACTTAACGTCACTCG 4079
Qy 1260 GlnLysAlaAlaAspValThrTyraAspSerGluValMetGlyValAlaGlyCysGlnGlu 1279
Db 4080 GAAAAAGTCACTGAAGTTGCCCTTAAAGGTGAAGGACAGAAAGAACCTGAATGTAAAG 4139
Qy 1280 LysGluSerThrGluValGlnSer-----LeuSerLeuGluGluGluGln 1294
Db 4140 GATGATGCTCTGTAAGTCAAGTCAAGTCTCCTCCATCCCGCTGGAGAGAGAG 4199
Qy 1295 MetGluThrAspValGluLysGluLysArgGluThrLysProGluGlnValSerGluGln 1314
Db 4200 ATGTAGATTCAAGTCGAAAGGAGAGAAACAGAAAGCAGAACCAACCATGTGATGAAGAG 4259
Qy 1315 Gly---GluGlnGluThrAlaAlaProGluHisGluGlyThrTyrlGlyLysProValLeu 1333
Db 4260 AAGCTTGACACAGAAACACTGTATCTGTAAGAGGTCAAGTGAAGCTCTCCAG 4319
Qy 1334 ThrLeuAspMetProSerSerGluArgGlyLysAlaLeuGlySerLeuGlySer--- 1352
Db 4320 ACGAGTAATGTCCCATCTAATATGGGCAAGAGAGTCAAGAGTTTGAAGAAACCTT 4379
Qy 1353 ProSerLeuProAspGlnAspLysAlaGlyCysIleGluValGlnValGlnSerLeuAsp 1372
Db 4380 CCTCCCTGCTAGTCAAGAGAGGAGGACGATATGACCAAAATTCAGTCTGAG 4439
Qy 1373 ThrThrValThrGlnThrAlaGluAlaValGluLysValIle-----GluThrValVal 1390

```

```

Db 4440 GCATCATTCACCTTAACGCGGCTGCAGAGAGAGAAAAAGTCTTAAGAGAACTGCCAAC 4499
Qy 1391 ILeSerGluThrGlyGluSerProGluCysValGlyAlaHisLeuLeuProAlaGluLys 1410
Db 4500 ATTTTAGAAACAGAGTGAACCTTGAGCCTCAGGTGACATCTTAAGTTCTGGAAGAGAA 4559
Qy 1411 SerSerAlaThrGlyGlyHisTrpThrLeuGlnHisAlaGluAspThrValProLeuGly 1430
Db 4560 TCTCTGAAAAAATGAAGACTTGGCCCTCATCCAGGGGAAAGTGTGTGCCACAGGG 4619
Qy 1431 ProGluSerGlnAlaGluSerIleProIleValThrProAlaProGluSerThrLeu 1450
Db 4620 CCGCAGCTGCAGGCAAAATTCACACAGTATGATATCTGTACTACCAAGAAAGGCTTA 4679
Qy 1451 HisProAspLeuGlnGlyGluIleSerAlaSerGlnArgGluArgSerGluGluLysP 1470
Db 4680 AGTTCCGACTCGAAGGAGAGAAAAACATCATCAGTAAGTGAAGTCAAGTCAAGTCAAT 4739
Qy 1471 LysProAspAlaGlyProAspAlaAspGlyLysGluSerThrAlaIleGluLysValLeu 1490
Db 4740 GAGCAGGTGCTTGCACAGAGGTC-----AAAGTGAAGTGAACATTTGAGAGGATTTA 4793
Qy 1491 LysAlaGluProGluIleLeuGluLeuGluSerLysSerAsnLysIleValLeuAsnVal 1510
Db 4794 GAGCTGAAATGGATTTTGAACCTTGAACACCAAAAGCAATTAATCTTCCAAACATC 4853
Qy 1511 ILeGlnThrAlaValAspGlnPheAlaArgThr---GluThrAlaProGluThrHisAla 1529
Db 4854 ATCCAGACAGCCGTGACCACTTGTATGATACAGAAAGAAAGCCAGCAAGATGTTGACG 4913
Qy 1530 TyraAspSerGlnThrGlnValProAlaCysArgLeuAspSerArgGluProAsnArgCys 1549
Db 4914 TCTGATTACAGACACCAAGCTCAGCTGATTAAGCTGACCCAGGAGCGTGGAGGAA 4973
Qy 1550 TrpThrLys-----MetLysAspAla-LysMetLysHisProValProGlu 1564
Db 4974 ACGAGAAAGAGAGAGAGAACTCAGGCTCTGACAGATGAAACCAATTAATCTTCA 5033
Qy 1564 nProArgGlu-----AspLeuGlnValLeuThrValLeuGluAl 1577
Db 5034 GCCAAAGAGAGTCAAGACTCAACGCAAGTGGAGACAGCACTTGTGATATTTCCAAAGAC 5093
Qy 1577 aTrpAlaGlnProArgLysCysLeuProArgLeuGlnLeuLysAlaProVal 1594
Db 5094 ATGAGTGAAGCTCAGAAAGACCATGACTGTGAGTGAAGGTTCCACTG 5145

RESULT 3
US-09-724-676A-12735
; Sequence 12735, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724, 676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12735
; LENGTH: 5931
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676A-12735

Alignment Scores:
Pred. No.: 3,65e-172 Length: 5931
Score: 4341.50 Matches: 970
Percent Similarity: 69.13% Conservative: 190
Best Local Similarity: 57.81% Mismatches: 408
Query Match: 53.78% Indels: 110
DB: 5 Gaps: 33

US-09-902-432-4 (1-1596) x US-09-724-676A-12735 (1-5931)

```


Dh 2307 GCATGGAGGAGGACCACAGAAAGCTGATGAGCCCGGAAAAAGACAAGAGAGCGGGGACA 2366
Qy 713 AspaLaValProAlaSerThrGlnGlnGlnAspGlnAlaGlnGlySerSerProGlu 732
Dh 2367 GACGGGATCTCTGCTGCTCCCAAGAACATGATCCAGGCGAGGAAAGTTCTCCCGGAG 2426
Qy 733 ProAlaGlySerProSerGlnGlnGlnValSerThrThrGlnSerPheLysArgLeu 752
Dh 2427 CAAGCTGGAAAGCTTACCAAGAGGGAGGGCGTTCCACTGGGAGTCATTAAAGGTTA 2486
Qy 753 ValThrProArgLysLysSerLysSerLysLeuGlnGlnLysAlaGlnAsp----- 769
Dh 2487 GTACAGCCCAAGAAAAAATCAAGTCCAGCTGGAAGAAAAAGCAAGACCTCCATAGCT 2546
Qy 770 ---SerSerValGlnGlnLeuSerThrGlnGlnProSerArgGlnGlnSerThrVal 788
Dh 2547 GGGCTGGGTGAACATTTCCACTCCAGCACTGAACCCGCTAAAGAAATCTGGGCTC 2606
Qy 789 SerLysLysPheLysPheProGlnArgLysLysArgAlaAspGlnLysGlnGln 808
Dh 2607 TCAATCAAGAAAGTTATTTCTGAGACGAAGAAAGGCGCAGATGGGAAACAGAAACA 2666
Qy 809 AlaThrValAlaAspSerGlyProValGlnLysGlnLysAspProAsnValProAla 828
Dh 2667 GCCCTGTTGAAGACCGGAGCCCAAGGGGCCCAAGAAAGTGACTGTATGTCCTCGGCTC 2726
Qy 829 ValValProLeuSerGlnLysAsnAlaValGlnArgGlnLysMetGlu-----Ala 845
Dh 2727 GTGGTCCCTCTGTGATGATGCTGTAGAAAGGGAAGAAATGAGGACACAGCAAGCC 2786
Qy 846 GlnGlnAsnThrGlnLeuProGlnLeuGlnValAlaValSerGlnLysLeuSer 865
Dh 2787 CAAAAAGCGCAGAGAGCGCCAGAGCAAGAGCACTGAGGTGTCAGAAAGGCTCACAC 2846
Qy 866 LysThrLeuValHisThrAlaSerValAlaValLysGlyThrArgAlaValThrSer 885
Dh 2847 GAGAGTCAGATTATATGATGAGAGAGCTGTCGTCGAGGAGCGAGGCGAGCTTACATT 2906
Qy 886 ValGlnGlnArgSerProSerThrLysAlaSerValThrGlnProLeuGlnHisThr 905
Dh 2907 ATTGAAGAAAGGCTCTCTTCTGATATCTGCTTCAGTACAGAACCTCTTGAACAAGTA 2966
Qy 906 AlaGlnGlnAlaMetProProValGlnGlnValThrGlnLysAspLysLeuAla---Gln 924
Dh 2967 GAAGCTGAAGCGCGCACTGTTAACTGAGAGGATTTGGAAAGAGAACTTAATGAGAAAG 3026
Qy 925 GlnThrProValLeuThrGlnThrLeuProGlnGlnLysAspAlaHisAspAspMetVal 944
Dh 3027 GAACCCCGCAGGTTATCTGAACCTCTGCCAGAAACAGAGAGGCGCGGGCGACACGTC 3086
Qy 945 ThrSerGlnValAspPheThrSerGlnAlaValThrAlaThrGlnThrSerGlnAlaLeu 964
Dh 3087 GTTAGAGAGCGGAATTGACCCCGGAAGCTGTGACAGCTGACAGAACTCAGGGCCATTG 3146
Qy 965 ArgThrGlnGlnValThrGlnAlaSerGlnAlaGlnGlnThrThrAspMetValSerAla 984
Dh 3147 GGTCGCGAAGAAAGAAACCAAGACATCTGCTGGAAGAGACCAACAATAATGGTGTAGCA 3206
Qy 985 ValSerGlnLeuThrAspSerProAspThrThrGlnGlnAlaThrProValGlnGlnVal 1004
Dh 3207 GTCTCCCAATTAAACGACTCCCAAGACACCAAGAGAGGCCACTCCGGTGCAGAGGTG 3266
Qy 1005 GlnSerGlnValLeuAspThrGlnGlnGlnGlnArgGlnThrGlnAlaLeuGlnAla 1024
Dh 3267 GAAGGTGGGCTACCTGACATAGAAAGAGCAAGAGCGGAGCTCAAGAGTCTCCAGGCA 3326
Qy 1025 ValAlaAspLysValLysGlnGlnGlnValProAlaThr----- 1038
Dh 3327 GTGGCGAAGAAAGTGAAGAGAAATCCAGCTGCGTGGACCGGTGGCCAGAAAGATGTG 3386
Qy 1039 ---GlnThrValGlnArgThrGlnSerLysAlaLeuGlnLysValGlnGlnValGlnGln 1057
Dh 3387 CTTCAAGCTGTGTGAGAGAGCA-----GAGGCAAGAAAGACCAAGAAAGAG 3428

Qy 1058 AspSerGlnValLeuAlaSerGlnLysGlnLysAspValMetProLysGlyProValGln 1077
Dh 3429 CAGGCTGAAGCGTCGGGTCTGAGAAAGAGACGAGTTGATGTTGAAGATGATGCTCAG 3488
Qy 1078 GlnAlaGlnAlaGlnHisLeuAlaGlnGlnLysSerGlnThrGlnGlnAlaThrProGlnSer 1097
Dh 3489 GAGCGAAAAACTGAGCCCTTTACACAAAGAGGTGTGGGGGACAGACCCCGCAAGAAAGC 3548
Qy 1098 LeuGln---ValProGlnValThrAlaAspValAspHis-----ValAlaThrCys 1113
Dh 3549 TTGGAAGAAAGCTCTCAAGTACAGAGAGCATAGAGTCACTGAGCTTGAACCACTTGT 3608
Qy 1114 GlnVal-----LysLeuGlnGlnLeu---MetGlnGlnAlaValAla 1127
Dh 3609 CAAGCCGAACCTTAGCTGGGGTAAATCAAGAGAGATGTGGAATGGAACAGGCTATCCCC 3668
Qy 1128 ProGlnSerSerGlnThrLeuThrAspSerGlnThrAsnGlySerThrProLeuAlaAsp 1147
Dh 3669 CCTGACTCGTGGAAACCCCTTACAGACAGTGAAGTGAATGAGACACCCCGTACCCGAC 3728
Qy 1148 SerAspThrAlaAspGlyThrGlnGlnAspGlnThrLysAspSerGlnAspSerLysAla 1167
Dh 3729 TTTGACGACACAGCAACAACCAAGAGAGAGATTTGGAATTCATGAGAGAGATGAG 3788
Qy 1168 ThrAlaAlaValArgGlnSerGlnValThrGlnGlnGlnAlaAlaThrAlaGlnLysGln 1187
Dh 3789 GTGGATCTGTGATCCCACTCAGGGGGCAGAGAGCGAGGGAGTTCTTGCACAGAAAG 3848
Qy 1188 GlnProSerThrLeuProAsnAsnValProAlaGlnGlnGlnHisGlnGlnLysProGly 1207
Dh 3849 AGGCTCCAGCA---CTTCCAGTTTGTGTTCCAGAAAGAACTAAAGAAACAATCAAG 3905
Qy 1208 ---ArgAspValLeuGlnProThrGlnGlnGlnLeuThrAlaAlaAlaValProValLeu 1226
Dh 3906 ATGGAAGACACTCTGAGACATATAGATTAAGAGGTGCTGAGTGAATCTGATTCCTTG 3965
Qy 1227 AlaLysThrGlnValGlnGlnGlnGlnGlnValAspThrPheLysAspGlyGlnLysValLys 1246
Dh 3966 TCMAAGACTGAGGGGAGCTCA---GAGGCTGACAGTATGCTGATGAGAAACCAAA 4019
Qy 1247 -----GlnGlnGlnGlnValPheValHisSerGly-----ProAsnSer 1259
Dh 4020 GACGTAACATTTTGAAGAGCTTAGGGGTCTATAGACACAGGATATAACACTGCTGG 4079
Qy 1260 GlnLysAlaAlaAspValThrThrAspSerGlnValMetGlnValAlaGlyCysGlnGln 1279
Dh 4080 GAAAGGTCACTGAAGTTGGCCCTTAAGGTGAAGAGAGAGCAAGAAAGCTGAATGTAAAG 4139
Qy 1280 LysGlnSerThrGlnValGlnSer-----LeuSerLeuGlnGlnGlnGln 1294
Dh 4140 GATGATGCTCTTGAACCTGAGAGTCACTTAAGTCTCTCCATTCCTCCGTGGAGAGAG 4199
Qy 1295 MetGlnThrAspValGlnLysGlnLysArgGlnThrLysProGlnGlnValSerGlnGln 1314
Dh 4200 ATGTGATTTCAAGTGAAGAGAGAAACAGAGAGAGACCAACCATGTGAATGAAGAG 4259
Qy 1315 Gln---GlnGlnGlnThrAlaAlaProGlnHisGlnGlnLysGlnLysThrGlnLysProValLeu 1333
Dh 4260 AAGCTTGAGACGAAGAACAGCTGTACCTGATCTGAAGAGGTGATGAACAGCTCTCCAG 4319
Qy 1334 ThrLeuAspMetProSerSerGlnArgGlnLysAlaLeuGlnSerLeuGlnGlnGln 1352
Dh 4320 ACAGTGAATGTGCCCATATAGATGGGCAAGAGAGCTGACAGTTTGAAGAGAGCCCT 4379
Qy 1353 ProSerLeuProAspGlnAspLysAlaGlyCysLysGlnGlnGlnValGlnGlnSerLeuAsp 1372
Dh 4380 CTTCCCTGCTGATGCTCAAGAGAGGAGCAAGTATGACCAATAATTCAGAGTCTGTAG 4439
Qy 1373 ThrThrValThrGlnThrAlaGlnAlaValGlnLysValLys-----GlnThrValVal 1390
Dh 4440 GCATCATTTCACTTAACAGCGGCTGCAGAGAGAGAAAGGTCTTGAAGAAATCTGCACAA 4499

Db 1011 AAGAAGCTGACAGATCTCCGACTAGTCCCGTACACGAGAAACAGATCA 1070
 QY SerPheLysLysPhePheThrHisGlyTTPAlaGlyTTPArgLysLysThrSerPheLys 302
 Db 1071 ACCCTCAAAAATTTCTTCACTCAAGGTTGGCGCGCTGGCGCAAAAAGACCAAGTTTCAGG 1130
 QY LysSerLysGluAspAspLeuGluThrAlaGluLysArgLysGluGluGluGluGluGlu 322
 Db 1131 AAGCCGAAGAGAGATGAACTGGAAGCTTCAAGAGAAAGAAAGAAACAGAACGACAGAAAAA 1190
 QY ValAspGluGluGluLysGluLysThrGluProAlaSerGluGlu----- 337
 Db 1191 GTAGACACGAAGAAAGCAAGAAAGCAAGCTTCCGAGAAACCTGACCCGCTCCGAG 1250
 QY 338 -----GlnGluProAlaGluAspThrAspGlnAlaArgLysSerAlaAspTyr 353
 Db 1251 CAAGCCACCCACAGAGAGCCGGCAAGAAAGTCCACAGAGCCCGGTTATCAGCTGAATAT 1310
 QY 354 GlnLysValGluLeuProLeuGluAspGlnValGlyAspLeuGluLysSerGluGlu 373
 Db 1311 GAGAAAGTTGAGCTGCCCTCAGAGAGCAAGTCAGTGGCTCCGAGGACCTTCTGAAGAG 1370
 QY 374 LysCysAlaProLeuAlaThrGluValPheAspGluLysMetGluAlaHisGln--Glu 392
 Db 1371 AAACCTGCTCCGTTGGCGACAGAGCTTTGATGAGAAATATGAACTCCACCAAGAAAG 1430
 QY 393 ValValAlaGluValHisValSerThrValGluLysThrGluGluGluGluGluGluGly 412
 Db 1431 GTTGTGGCGAAGTCCAGCTCAGACCCGTGAGAGAGAAACGAAAGAGCAG----- 1481
 QY 413 GlyGluAlaGluGluGlyValValValGluGluGlyThrGlyGluSerLeuProGluLys 432
 Db 1482 -----AAAAAGAGGTGGAAAGAAACAGAGGCTCTTGTCCAGCTGAAGAA 1526
 QY 433 LeuAlaGluProGluGluValProGluGluAlaGluProAlaGluGluLeuMetLysSer 452
 Db 1527 TTGGTTGGAATGATGCAGAACTCAGAAAGCCGAACCTGCCAAGAGAGCTGGTGAAGCTC 1586
 QY 453 ArgGluMetCysValSerGlyLysAspHisThrGlnLeuThrAspLeuSerProGluGlu 472
 Db 1587 AAGAAACGTGTGTTCGCGAGAGAGACCTTACAGAGAGAGCTCAGCTCATGTAGAG 1646
 QY 473 LysThrLeuProLysHisProGluGluGlyLeuValSerGluValGluMetLeuSerGln 492
 Db 1647 AAGGTCTCTCCAAACCCCGAAGGCGTTGTGAGGTGGAATGCTGTATACAG 1706
 QY 493 GlnArgLysLeuValGlnGluSerProLeuLysLysPheSerSerSerGlyLeuLys 512
 Db 1707 GAGAGAAATGAAGTGCAGGGAAGTCCATTAAGAAAGCTTTTACCAGCAGCTGCTTAAA 1766
 QY 513 LysLeuSerGlyLysLysGluLysGlyLysArgGlyGlyGlyLysAspGluGluProGly 532
 Db 1767 AAGCTTTCTGGAAGAAAGAAACGAAAGGAAAAAG--GAGAGAGACAGAGAAATCAGGG 1823
 QY 533 GlnTyrGlnHisIleHisThrGluSerProGluSerAlaAspGluGluLysGlyLysSer 552
 Db 1824 GAGCACTACCTGATCCAGCCGATTTCTCCGACAGCCAGAGAGAGAAAAAGGCGAGAC 1883
 QY 553 SerAlaSerSerProGluGluProGluGluThrThrCysLeuGluLysGlyProLeuGlu 572
 Db 1884 TCTGCTTATCCCTCTAGAGAGCCGAGAGAGATACGTGTCTGAAAAAGGCTTATGCCGAG 1943
 QY 573 AlaProGlnAspGlyGluAlaGluGluGlyThrThrSerAspGlyGluLysLysArgGlu 592
 Db 1944 GAGCAGACAGAGATGGAGAGCTGAAGAGAGCTACTTCGATGAGAAAAAGAAAAAGAGAA 2003
 QY 593 GlyIleThrProTTPAlaSerPheLysLysMetValThrProLysLysArgValaArg 612
 Db 2004 GGTTGACTCTCCCTGGGCAATTCATTAAGAAAGATGTGACGCCCAAGAAAGGTGTAAACGG 2063
 QY 613 ProSerGluSerAspLysGluGluGluGluGluValLysSerAlaThrLeuSerSer 632
 Db 2064 CTTTCGAAAGATGATTAAGAAAGATGAGCTGGAACAGGTCAAGAGCCTTCTGTCTTCC 2123

QY 633 ThrAspSerThrValSerGluMetGlnAspGluValLysThrValGlyGluGluLys 652
 Db 2124 ACCGAGACACAGAGCTCTGGAATGTCAGAGAAAGAAAGAGAGCGGTGAAGAGCAAG 2183
 QY 653 ProGluGluProLysArgArgValAspThrSerValSerTTPGluAlaLeuIleCysVal 672
 Db 2184 CCGAGAAACCAAGCCGAGGTGATACCTCAGTATCTTGGAAAGCTTAAATTTGTGTG 2243
 QY 673 GlySerSerLysLysArgAlaArgLysAlaSerSerSerAspAspGluGlyProArg 692
 Db 2244 GGATCATCAAGAAAGAGCAAGAGAGAGTCTCTTCTGATGAGAGAGGGGAGCAAAA 2303
 QY 693 ThrLeuGlyGlyAspSerHisArgAlaGluGluAlaSerLysAspLysGluAlaGlyThr 712
 Db 2304 GCATGGAGAGAGACACACAGAAAGCTGATAGCGCCGAGAAAGCAAGAGAGCGGAGCA 2363
 QY 713 AspAlaValProAlaSerThrGlnGluGluGlnAspGlnAlaGlnGlySerSerProGlu 732
 Db 2364 GAGGGATCTCTTGTCTGTTCCAGAAACATGATCCAGGGAGAGGAGAGTTCTCCCGAG 2423
 QY 733 ProAlaGlySerProSerGluGlyGluGlyValSerThrTTPGluSerPheLysArgLeu 752
 Db 2424 CAAGCTGGAGAGCTTACCGAAGAGGAGGCGTTTCCACTCGGAGTCAATTAAAGCTTA 2483
 QY 753 ValThrProArgLysLysSerLysSerLysLeuGluGluValAlaGluAsp----- 769
 Db 2484 GTACGCGCAAGAAAGAAATCAAGTCCAAAGCTGGAAGAGAAAGCAAGACCTCATAGCT 2543
 QY 770 ---SerSerValGluGluLeuSerThrGluIleGluProSerArgGluGluSerThrVal 788
 Db 2544 GGCTGTGTTGTAACATTCCACTCAGACACTGAACCCGGTAAAGAAATCTGGGTC 2603
 QY 788 SerIleLysPheIleProGlyArgArgLysLysArgAlaAspGlyLysGluGluGln 808
 Db 2604 TCAATCAAGAAAGTTATTCCTGACGAGAAAGAAAGAGCCAGATGGGAAACCAAGACA 2663
 QY 809 AlaThrValGluAspSerGlyProValGluIleAsnGluAspAspProAsnValProAla 828
 Db 2664 GCCCTGTGTAAGAGAGCGAGGCGCAAGAGGCGCAAGAAATGATCTGATCTCCGGCC 2723
 QY 828 ValValProLeuSerGlyTyrAsnAlaValGluArgGluLysMetGlu-----Ala 845
 Db 2724 GTGTCTCTCTCTGAGTATGATCTGTAAGAAAGGAAATGAGACACAGAGCC 2783
 QY 846 GlnGlyAsnThrGluLeuProGluLeuGluGlyAlaValTyrValSerGluGluLeuSer 865
 Db 2784 CAAAAGGCGCAGACACCCGAGCAGAAAGCAGCAGCTGAGGTCCAGAGAGCTCAGC 2843
 QY 866 LysThrLeuValHisThrValSerValAlaValIleAspGlyThrArgAlaValThrSer 885
 Db 2844 GAGATCAGGTTCATATGATGTCAGACACAGCTGCTGTCAGGACAGAGGCGAGCTTCCATT 2903
 QY 886 ValGluGluArgSerProSerThrIleSerAlaSerValThrGluProLeuGlnHisThr 905
 Db 2904 ATTGAAGAAAGGTCTCTCTTGGATATCTGCTTATGAGACAGAACTCTTGAACAAGTA 2963
 QY 906 AlaGlyGluAlaMetProProValGluGluValThrGluLysAspIleIleAla--Glu 924
 Db 2964 GAAGCTGAAGCGCAGCTGTAACTGAGAGGTAATTGAAGAAAGAGATTAATTCAGAGAA 3023
 QY 925 GlnThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHisAspMetVal 944
 Db 3024 GAACCCCGCAGCTTACGTGAACCTCTCCAGAGAAACAGAGAGCCCGGCGCAGCAGCGTTC 3083
 QY 945 ThrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeu 964
 Db 3084 GTTATGAGGCGGAATTGACCCCGAAGCTGTGACAGCTGCAGAACTGCAGAGGCCATTG 3143
 QY 965 ArgThrGluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSerAla 984
 Db 3144 GGTTCGAAAGAGAAACGAGAGCATCTGCTGTGAAGAGACCAAGAAATGTTGTACAGA 3203

ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (91)..(91)
 OTHER INFORMATION: n is a,c,g, or t
 US-09-724-676-12736

Alignment Scores:

Pred. No.:	2,33e-162	Length:	5773
Score:	4104.50	Matches:	912
Percent Similarity:	69.40%	Conservative:	179
Best Local Similarity:	58.02%	Mismatches:	389
Query Match:	50.84%	Indels:	92
DB:	5	Gaps:	29

US-09-902-432-4 (1-1596) x US-09-724-676-12736 (1-5773)

```

QY 89 ValGlgIaArgGluSerGluAspValArgGluLysAspArgValGluGluMetAla1a 108
DB 352 GTTGGACAGAGAGACTCTGAAGATGTGAGCGAAGAGACTCCGATTAAGAGATGCTACT 411
QY 109 AsnSerThrAlaValGluAspIleThrLysAspGlyGlnGluGluThrSerGluIle1e 128
DB 412 AACTCAGCGGTTTTCACGACATCACAGATGATGGCAGAGAGACACCCGAATATATC 471
QY 129 GluGlnIleProAlaSerGluAsnAsnValGluGluMetValGlnProAlaGluSerGln 148
DB 472 GAACAAATTCCTCTTCCAGAAACCAATTAGAAAGCTAACACACCCCTGAGTCCAG 531
QY 149 AlaAsnAspValGlyPheLysLysValPheLysPheValGlyPheLysPheThrValLys 168
DB 532 GCTAAAGATATGGATTAAAGAGGTGTTAACTTTGTTGCTTTAAATTCACTGTGAA 591
QY 169 LysAspLysAsnGluLysSerAspThrValGlnLeuLeuThrValLysLysAspGluGly 188
DB 592 AAGGATTAAGACAGAGAGAGCCTGACCTGTCACTCACTCACTGAAAGAAAGTGAAGG 651
QY 189 GluGlyAlaGluLysAspValGlyAlaGlyAspHisGlnGluProSerValGluThrAla 208
DB 652 GAGGGAGCA-----GCAGGGGCTGGCGACACACAGAGACCCAGCCCTT-----GGG 696
QY 209 ValGlgIaSerLysLysSerGluSerGluLysGlnSerThrGluLysGlnGluGly 228
DB 697 GCTGGAGAGAGAGATCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 756
QY 229 ThrLeuLysGlnGluLysSerThrGluLysLeuProLeuGlnAlaGluSerAspGlnAla 248
DB 757 ACCCTGAAGCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 816
QY 249 AlaGluGluGluAlaLysAspGluGlyGluGluLysGlnGluLysGluProThrLysSer 268
DB 817 GTC---GAGGAATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 873
QY 269 ProGluSerProSerSerProValAsnSerGluThrThrSerSerPheLysLysPhePhe 288
DB 874 GCAGATCTCCGACTAGTCCGTCAGCAGTGAACAGAGATCACTTCAAAAAATCTTTC 933
QY 289 ThrHisGlyThrAlaGlyThrAlaGlyLysLysThrSerPheLysSerLysGluAspAsp 308
DB 934 ACTCAAGGTTGGCCCGCTGGCCCAAAAGACCAAGTTTCAGAAAGCCGAAGAGAGATGA 993
QY 309 LeuGluThrAlaGluLysArgLysGlnGlnGluAlaGluLysValAspGluGluLys 328
DB 994 GTGAGAGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1053
QY 329 GluLysThrGluProAlaSerGluGlu-----GlnGlu 339
DB 1054 GGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1113
QY 340 ProAlaGluAspThrAspGlnAlaGlyLeuSerAlaAspThrGluLysValGluLeuPro 359
DB 1114 CCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1173

```

```

QY 360 LeuGluAspGlnValGlyAspLeuGluLysSerSerGluGluLysCysAlaProLeuAla 379
DB 1174 TCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1233
QY 380 ThrGluValPheAspGluLysMetGluAlaHisGln---GluValAlaAlaGluValHis 398
DB 1234 ACAGAAAGCTTTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1293
QY 399 ValSerThrValGluLysThrGluGluGlnGluGlyGlyGlyGlyGluAlaGluGlyGly 418
DB 1294 GTCCACACCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1329
QY 419 ValValValGluGlyThrGlyGluSerLeuProGluLysLeuAlaGluProGlnGlu 438
DB 1330 ACGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1389
QY 439 ValProGlnGluAlaGluProAlaGluLysLeuMetLysSerArgGluMetCysValSer 458
DB 1390 GAACCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1449
QY 459 GlyLysAspHisThrGlnLeuThrAspLeuSerProGluGluLysThrLeuProLysHis 478
DB 1450 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1509
QY 479 ProGluGlyIleValSerGluValGluMetLeuSerSerGlnGluAlaGlyLysValGln 498
DB 1510 CCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1569
QY 499 GlySerProLeuLysLeuPheSerSerSerGlyLeuLysLysLeuSerGlyLysLys 518
DB 1570 GGAAGTCCACTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1629
QY 519 GlnLysGlyLysArgGlyGlyGlyGlyLysAspGluGluProGlyGluTyrglnHis 538
DB 1630 CAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1686
QY 539 ThrGluSerProGluSerAlaAspGluGlnLysGlyGluSerSerAlaSerSerProGlu 558
DB 1687 GCCGATCTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1746
QY 559 GluProGluGluThrThrCysLeuGluLysGlyProLeuGluAlaProGlnAspGlyGlu 578
DB 1747 GAGCCGAGAGAGATACGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1806
QY 579 AlaGluGluGlyThrThrSerAspGlyLysLysArgGluGlyIleThrProThrAla 598
DB 1807 GCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1866
QY 599 SerPheLysLysMetValThrProLysLysArgValArgProSerGluSerAspLys 618
DB 1867 TCATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1926
QY 619 GluGluGluLeuGluLysValLysSerAlaThrLeuSerSerThrAspSerThrValSer 638
DB 1927 GAAAGATGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1986
QY 639 GluMetGlnAspGluValLysThrValGlyGluGluGlnLysProGluGluProLysArg 658
DB 1987 GAAATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2046
QY 659 ArgValAspThrSerValSerThrProLysAlaLeuIleCysValGlySerSerLysArg 678
DB 2047 AAGGTGAGATCCTCACTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2106
QY 679 AlaArgLysAlaSerSerSerAspAspGluGlyGlyProArgThrLeuGlyLysAspSer 698
DB 2107 GCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2166
QY 699 HisArgAlaGluGluLysSerLysAspLysGluAlaGlyThrAspAlaValProAlaSer 718
DB 2167 CAGAAAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2226
QY 719 ThrGlnGluGlnAspGlnAlaGlnLysSerSerProGluProAlaGlySerProSer 738

```

Db	2227	TCCCAAGACATGATCCAGGCGAGGAAGTTCTCCCGGAGCAAGCTGGAAGCCCTTACC	2286
Qy	739	GlUGlyGluGlyValSerThrTrpGluSerPheLYsArgLeuValThrProArgLYsLYs	758
Db	2287	GAAGGGGAGGGCGTTTCCACCTGGAGTCA'TTAAAGGTTAGTCACGCCCAAGAAAAA	2346
Qy	759	SerLYsSerLYsLeuGluGluLYsAlaGluAsp-----SerSerValGluGln	774
Db	2347	TCAAAGTCCAAGCTGGAAGAGAAAAAGCAAGACTCCATAGCTGGTGTGTGTAGAACAT	2406
Qy	775	LeuSerThrGluIleGluProSerArgGluGluSerTrpValSerIleLYsLYsPheIle	794
Db	2407	TCCACTCCAGACTGAACCCGGTAAAGAAAGATCCTGGGTCTCAATCAAGAAGTTTATT	2466
Qy	795	ProGlyArgArgLYsLYsArgAlaAspGlyLYsGlnGluGlnAlaThrValGluAspSer	814
Db	2467	CCTGGACGAAGGAAGAAAGGCCAGATGGGAACAAGAACAGCCCTGTTGAAGAGCCA	2526
Qy	815	GlyProValGluIleLeuGlnAspAspProAsnValProAlaValValProLeuSerGlu	834
Db	2527	GGGCCAACAGGGGCCAACGAAGATGACTCTGATGTCCCGCGGTGCTCTGTCTGTGAG	2586
Qy	835	TyrAsnAlaValGluArgGluLYsMetGlu-----AlaGlnGlyAsnThrGluLeu	851
Db	2587	TATGATGCTGTAGAAGGGAGAAATGGAGGCACAGCAAGCCCAAAAGACGGAGCAG	2646
Qy	852	ProGlnLeuLeuGlyAlaValTyrValSerGluGluLeuSerLYsThrValHisThr	871
Db	2647	CCGAGCAGAAGCGACCCACTGAGGTGTCCAGGAGCTCAGCAGAGTCAAGTTCATATG	2706
Qy	872	ValSerValAlaValIleAspGlyThrArgAlaValThrSerValGluGluArgSerPro	891
Db	2707	ATGGCAGCAGCTCTCGCTGACGGGACGAGCGCAGTACCATTATTGAAGAAAGTCTCCT	2766
Qy	892	SerTrpIleSerAlaSerValThrGluProLeuGluHisThrAlaGlyGluAlaMetPro	911
Db	2767	TCTTGTGATATCTGCTTCAGTGCAGAACCTCTTTGAACAGTAGAAGCTGAAGCGGCATG	2826
Qy	912	ProValGluValThrGluLYsAspIleAla---GluGluThrProValLeuThr	930
Db	2827	TTAACTGAGGAGGTATTGGAAAGAGAAGTAATTGCAGAGAAGAACCCCCACGGTTACT	2886
Qy	931	GlnThrLeuProGluGlyLYsAspAlaHisAspAspMetValThrSerGluValAspPhe	950
Db	2887	GAACCTCTGCCAGAGAACAGAGAGGCCGGGGCGACACGGTCTGTAGTGAGGCGGAATTG	2946
Qy	951	ThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeuArgThrGluGluValThr	970
Db	2947	ACCCCCGAAGCTGTGACAGCTGCAGAAACTGCAGGGCCATTTGGGTGCCGAAGAAGCAAC	3006
Qy	971	GluAlaSerGlyAlaGluGluThrThrAspMetValSerAlaValSerGlnLeuThrAsp	990
Db	3007	GAAGCATCTGCTCTGAAGAGACCAACGAATAATGGTGTACAGACTCTCCCAAGTTAACCGAC	3066
Qy	991	SerProAspThrThrGluGluAlaThrProValGlnGluValGluSerGlyValLeuAsp	1010
Db	3067	TCCCCAGACACACAGAGGAGGCCACTCCCGTGCAGGAGGTGGAAGGTGCGTACCTGAC	3126
Qy	1011	ThrGluGluGluGluArgGlnThrGlnAlaIleLeuGlnAlaValAlaAspLYsValLYs	1030
Db	3127	ATAGAGAGCAAGAGAGGGCGGACTCAAGAGGTCTCCAGGCAGTGGCAGAAAAAGTGAAA	3186
Qy	1031	GluGluSerGlnValProAlaThr-----GlnThrValGlnArg	1043
Db	3187	GAGAAATCCAGCTGCTTCGACCTGGTGGGCCAGAGATGTGCTTACGCTGTGCAGAGA	3246
Qy	1044	ThrGlySerLYsAlaLeuGluLYsValGluValGluGluAspSerGluValLeuAla	1063
Db	3247	GCA-----GAGGCAGAAAGACCAAGAGCAGGAGTGAAGCGTGGGT	3288
Qy	1064	SerGluLYsGluLYsAspValMetProLYsGlyProValGlnGluAlaGlyAlaGluHis	1083

3289	CTGAAGAAAGAGACCGGATGTAGTGTGTTAAAGGTAGATGCTCAGGAGGCCAAAACCTGAGCCT	3344	CTGAAGAAAGAGACCGGATGTAGTGTGTTAAAGGTAGATGCTCAGGAGGCCAAAACCTGAGCCT
QY	LeuAlaGlnGlySerGluThrGlyGlnAlaThrProGluSerLeuGlu---ValProGlu	1102	LeuAlaGlnGlySerGluThrGlyGlnAlaThrProGluSerLeuGlu---ValProGlu
DB	TTTACACAAAGGAAGGTGGTGGGGCAGACCACCCCAAGAAAGCTTTGAAAAAGCTCCTCAA	3408	TTTACACAAAGGAAGGTGGTGGGGCAGACCACCCCAAGAAAGCTTTGAAAAAGCTCCTCAA
QY	ValThrAlaAspValAspHis-----ValAlaThrCysGlnVal-----	1115	ValThrAlaAspValAspHis-----ValAlaThrCysGlnVal-----
DB	GTACACAGAGACATAGAGTCCAGTGTGATGAGTGTGTAACACCACTGTGTAAGCCGCAACCTTAGCT	3468	GTACACAGAGACATAGAGTCCAGTGTGATGAGTGTGTAACACCACTGTGTAAGCCGCAACCTTAGCT
QY	--llellylsLeuGlnLeu---MetGluGlnAlaValAlaProGluSerSerGluThr	1133	--llellylsLeuGlnLeu---MetGluGlnAlaValAlaProGluSerSerGluThr
DB	GGGTAAATAATCACAGGAGATGTGTATGTAAGAGCTATCCCCCTGACTCGGTGGAAACC	3528	GGGTAAATAATCACAGGAGATGTGTATGTAAGAGCTATCCCCCTGACTCGGTGGAAACC
QY	LeuThrAspSerGluThrAsnGlySerThrProLeuAlaAspSerAspThrAlaAspGly	1153	LeuThrAspSerGluThrAsnGlySerThrProLeuAlaAspSerAspThrAlaAspGly
DB	CCTACAGACACTGAGACTGTATGAAGCACCACCCGTAGCCGACTTTGACGCACACCGGCACA	3588	CCTACAGACACTGAGACTGTATGAAGCACCACCCGTAGCCGACTTTGACGCACACCGGCACA
QY	ThrGlnGlnAspGluThrIleAspSerGlnAspSerLysAlaThrAlaAlaValargGln	1173	ThrGlnGlnAspGluThrIleAspSerGlnAspSerLysAlaThrAlaAlaValargGln
DB	ACCCAGAAAGACGAGATTGTGGAATTCCTATGAGGAATGAGTGCATCTCGTATCCCGAC	3648	ACCCAGAAAGACGAGATTGTGGAATTCCTATGAGGAATGAGTGCATCTCGTATCCCGAC
QY	SerGlnValThrGluGluGluAlaAlaThrAlaGlnLysGluGluProSerThrLeuPro	1193	SerGlnValThrGluGluGluAlaAlaThrAlaGlnLysGluGluProSerThrLeuPro
DB	TCAGGGGCACAGAAGCAGAGGCAGTTCCTGCACAGAAAGAGAGCCCTCCACGA---CCT	3705	TCAGGGGCACAGAAGCAGAGGCAGTTCCTGCACAGAAAGAGAGCCCTCCACGA---CCT
QY	AsnAsnValProAlaGlnGluGluHisGlyGluGluProGly---ArgAspValLeuGlu	1212	AsnAsnValProAlaGlnGluGluHisGlyGluGluProGly---ArgAspValLeuGlu
DB	TCCAGTTTGTGTTCCAGGAAGAACTAAGAACAACATCAAGATCGAAGACACTCTAGAG	3765	TCCAGTTTGTGTTCCAGGAAGAACTAAGAACAACATCAAGATCGAAGACACTCTAGAG
QY	ProThrGlnGlnLeuLeuThrAlaAlaAlaValProValLeuAlaLysThrGluValGly	1232	ProThrGlnGlnLeuLeuThrAlaAlaAlaValProValLeuAlaLysThrGluValGly
DB	CATACAGATAAAGAGGTGTCTAGTGGAAACTGTATCCATCTCTCTCAAGACTGAGGGACT	3825	CATACAGATAAAGAGGTGTCTAGTGGAAACTGTATCCATCTCTCTCAAGACTGAGGGACT
QY	GlnGluGlyGluValAspTrpLeuAspGlyGluLysValLys-----Glu	1247	GlnGluGlyGluValAspTrpLeuAspGlyGluLysValLys-----Glu
DB	CAAA-----GAGCGTGACCATGCTGTATGTAGAGAAACCAAGACGTCACATTTTTCGAA	3879	CAAA-----GAGCGTGACCATGCTGTATGTAGAGAAACCAAGACGTCACATTTTTCGAA
QY	GluGlnGluValPheValHisSerGly-----ProAsnSerGlnLysAlaAlaAspVal	1265	GluGlnGluValPheValHisSerGly-----ProAsnSerGlnLysAlaAlaAspVal
DB	GGACTTGAGGGTCTATAGACACAGGCATAACNGTCAGTCGGGAAAGGTCACTGAAGTT	3939	GGACTTGAGGGTCTATAGACACAGGCATAACNGTCAGTCGGGAAAGGTCACTGAAGTT
QY	ThrTyrAspSerGluValMetGlyValAlaGlyCysGlnGluLysGluSerThrGluVal	1285	ThrTyrAspSerGluValMetGlyValAlaGlyCysGlnGluLysGluSerThrGluVal
DB	GCCTTTAAAGGTGAAGGCACAGAAGAAGCTGAATGTAAAAAGGATGATGCTCTTGAACTG	3999	GCCTTTAAAGGTGAAGGCACAGAAGAAGCTGAATGTAAAAAGGATGATGCTCTTGAACTG
QY	GlnSer-----LeuSerLeuGluGluGlyGluMetGluThrAspValGlu	1300	GlnSer-----LeuSerLeuGluGluGlyGluMetGluThrAspValGlu
DB	CAGAGTACCGTAACTCTCCATCCCGTGGAGAGAGATGGTAGTTCAGTTCGAA	4059	CAGAGTACCGTAACTCTCCATCCCGTGGAGAGAGATGGTAGTTCAGTTCGAA
QY	LysGluLysArgGluThrLysProGluGlnValSerGluGluGly---GluGlnGluThr	1319	LysGluLysArgGluThrLysProGluGlnValSerGluGluGly---GluGlnGluThr
DB	AGGGNGAAAACAGAACAGACAGCCCAACCCATGTGAATGAAGAGAAGCTTGAGCAGCAACA	4119	AGGGNGAAAACAGAACAGACAGCCCAACCCATGTGAATGAAGAGAAGCTTGAGCAGCAACA
QY	AlaAlaProGluHisGluGlyThrTyrGlyLysProValLeuThrLeuAspMetProSer	1339	AlaAlaProGluHisGluGlyThrTyrGlyLysProValLeuThrLeuAspMetProSer
DB	GCTGTTACCGTATCTGAAGAGGTGAGTAAGAGCTCCTCCACACAGTGAATGTGCCCATC	4179	GCTGTTACCGTATCTGAAGAGGTGAGTAAGAGCTCCTCCACACAGTGAATGTGCCCATC
QY	SerGluArgGlyLysAlaLeuGlySerLeuGlyGlySer---ProSerLeuProAspGln	1358	SerGluArgGlyLysAlaLeuGlySerLeuGlyGlySer---ProSerLeuProAspGln
DB	ATAGTGGGGCAAGGAAGTCAGCAGCTTTTGGAAAGGAGCCCTCTCCCTGCTAGGTCAA	4239	ATAGTGGGGCAAGGAAGTCAGCAGCTTTTGGAAAGGAGCCCTCTCCCTGCTAGGTCAA
QY	AspLysAlaGlyCysIleGluValGlnValGlnSerLeuAspThrThrValThrGlnThr	1378	AspLysAlaGlyCysIleGluValGlnValGlnSerLeuAspThrThrValThrGlnThr
DB	GAGGAGCATATGCACCAAAATTCAGTTTCAGAGCTCTGAGGCATCATCTACTCTAACA	4299	GAGGAGCATATGCACCAAAATTCAGTTTCAGAGCTCTGAGGCATCATCTACTCTAACA
QY	AlaGluAlaValGluLysValIle-----GluThrValValIleSerGluThrGlyGlu	1396	AlaGluAlaValGluLysValIle-----GluThrValValIleSerGluThrGlyGlu
DB	CGCGCTGCAGAGGAGGAAAAGGCTTTAGAGGAAACTGCCCAACATTTTAGAACAGGTGAA	4359	CGCGCTGCAGAGGAGGAAAAGGCTTTAGAGGAAACTGCCCAACATTTTAGAACAGGTGAA
QY	SerProGluCysValGlyAlaHisLeuLeuProAlaGluLysSerSerAlaThrGlyGly	1416	SerProGluCysValGlyAlaHisLeuLeuProAlaGluLysSerSerAlaThrGlyGly
DB	ACGTTGAGCCCTGCAGGTGCACATTTAGTTCTGGAAGAGAAATCTCTCAAAAAAATGAA	4419	ACGTTGAGCCCTGCAGGTGCACATTTAGTTCTGGAAGAGAAATCTCTCAAAAAAATGAA

D	b	1390	GAACTCAGGAGCCGNACTGCCAAGGAGCTGGTGAAGCTCAAGAAAACGTGTGTTTCC	1449
Q	y	459	GlyGlyAspHisThrGlnLeuThrAspLeuSerProGluGluLysThrLeuProLysHis	478
D	b	1450	GGAGGAGACCTACACAGGAGCTGACCTCAGTCTGATGAGAAGTGCTGTCTCCAAACCC	1509
Q	y	479	ProGluGlyIleValSerGluValGluMetLeuSerSerGlnGluArgIleLysValGln	498
D	b	1510	CCCGAAGCGCTTGTGAGTGAGTGGGAAATGCTGTCTATCACAGGAGAAATCAAGGTGCAG	1569
Q	y	499	GlySerProLeuLysLysLeuPheSerSerGlyLeuLysLysLeuSerGlyLysLys	518
D	b	1570	GGAGTCCACTAAGAAGCTTTTACCACACTGCTTAAAAAGCTTCTGGAAAGAA	1629
Q	y	519	GlnLysGlyLysArgGlyGlyGlyAspGluGluProGlyGluThrGlnHisIleHis	538
D	b	1630	CAGAAAGGAAAAAGAA--GGAGGAGAGACGAGGAATCAGGGGAGCACTCAGGTCCCA	1686
Q	y	539	ThrGluSerProGluSerAlaAspGluGlnLysGlyGluSerSerAlaSerSerProGlu	558
D	b	1687	GCCGATTCTCCGACACCGCAGGAGGAGCAAAAGGCGGAGAGCTCTGCCTCATCCCTCGAG	1746
Q	y	559	GluProGluGluThrThrCysLeuGluLysGlyProLeuGluAlaProGlnAspGlyGlu	578
D	b	1747	GAGCCCGAGGAGATCAGTGTCTGGAAAAGGCTTAGCCGAGGTGCAGAGATGGGGAA	1806
Q	y	579	AlaGluGluGlyThrThrSerAspGlyGluLysLysArgGluGlyIleThrProTrpAla	598
D	b	1807	GCTGAAGAGGAGCTACTTCCGATGAGAGAAAAAAGAGAAGGTGTCACTCCCTGGGCA	1866
Q	y	599	SerPheLysLysMetValThrProLysLysArgValArgProSerGluSerAspLys	618
D	b	1867	TCATTCAAAAGATGGTGCAGCCCAAGAGCGTGTAGACGCGCTTCGAAAGTGATAAA	1926
Q	y	619	GluGluGluLeuLysValLysSerAlaThrLeuSerSerThrAspSerThrValSer	638
D	b	1927	GAAATGAGCTGCACAAAGGTCAAGAGCGTACTCTGTCTTCCACGAGAGCACAGCTCT	1986
Q	y	639	GluMetGlnAspGluValLysThrValGlyGluGluGlnLysProGluGluProLysArg	658
D	b	1987	GAAATGCAAGAGAAATGAAGGAGCGCTGGAAAGAGCCANAGCCGGAAGAACCAAGCGC	2046
Q	y	659	ArgValAspThrSerValSerTrpGluAlaLeuIleCysValGlySerSerLysLysArg	678
D	b	2047	AAGTGGATACCTCAGTATCTTGGGAAGCTTAAATTTGTGTGGATCATCCAGAAAAGA	2106
Q	y	679	AlaArgLysAlaSerSerSerAspAspGluGlyGlyProArgThrLeuGlyGlyAspSer	698
D	b	2107	GCAAGGAGGGTCCCTCTCTGTATGGAAGGGGGACCAAAAGCAATGGGAGGAGCACAC	2166
Q	y	699	HisArgAlaGluAlaSerLysAspLysGluAlaGlyThrAspAlaValProAlaSer	718
D	b	2167	CAGAAAGCTGATGAGCGCGGAAAGACAAAGAGACGGGACAGCGGATCCTTGCTGGT	2226
Q	y	719	ThrGlnGluGlnAspGlnAlaGlnGlySerSerSerProGluProAlaGlySerProSer	738
D	b	2227	TCCCAAGAACATGATCCAGCGCAGGGAAGTCTCTCCCGGAGCAAGCTGGAAGCCCTACC	2286
Q	y	739	GluGlyGluGlyValSerThrTrpGluSerPheLysArgLeuValThrProArgLysLys	758
D	b	2287	GAAGGGAGGGCGTTTCCACTCGGAGTCATTTTAAAGGTTAGTCACGCCAAGAAAAAAA	2346
Q	y	759	SerLysSerLysLeuGluGluLysAlaGluAsp-----SerSerValGluGln	774
D	b	2347	TCAAAGTCCAAGCTTGTGAAGAGAAACGGAAGACTCCATAGCTGGGTCTGGTGAACAAT	2406
Q	y	775	LeuSerThrGluIleGluProSerArgGluGluSerTrpValSerIleLysLysPheIle	794
D	b	2407	TCCACTCCAGACACTGAACCCCGGTAAAGAAGATCCTGGGTCTCAATCAAGAAGTTATT	2466
Q	y	795	ProGlyArgArgLysLysArgAlaAspGlyLysGlnGluGlnAlaThrValGluAspSer	814
D	b	2467	CCTGGACGAAGGAAGAAAGCCGAGATGGGAAACAAAGAACCAAGCCCTGTTGAAGACGCA	2526

Qy	815	GlyProValGluIleuLenGluAspAppProAenValProAlaValValProIeuSerGlu	834
Db	2527	GGGGCCACAGGGGCCAACAGAGATGACTCTGATGTCGGCGCGGTGCTCCCTCTGCTGAG	2586
Qy	835	TyrAenAlaValGluArgGluIlysMetGlu-----AlaGlnGlyAsnThrGluLeu	851
Db	2587	TATGATGCTGTAGAAAGGGAGAAATATGGAGGCACAGCAAGCCCAAAAGGCGGAGCAG	2646
Qy	852	ProGlnLeuLeuGlyAlaValTyrValSerGluLeuLeuSerLysThrLeuValHisThr	871
Db	2647	CCCGAGCAGAAGCGCACCTAGTAGGTGTCCAAGSAGCTCAGCGAGAGTCAGGTTTCATATG	2706
Qy	872	ValSerValAlaValIleAspGlyThrArgAlaValThrSerValGluGluArgSerPro	891
Db	2707	ATGCGACAGCTGTGCTGTACGGCAGCGAGGAGCTACCATTTATGAAGAAAGGTCTCCT	2766
Qy	892	SerThrIleSerAlaSerValThrGluProLeuGluHisThrAlaGlyGluAlaMetPro	911
Db	2767	TCCTTGGATATCTGCTTCAGTGCAGAACTCTTGAACAGGTAGAAAGCTGAAGCCGCAC	2826
Qy	912	ProValGluGluValThrGluIlysAspIleAla---GluGluThrProValLeuThr	930
Db	2827	TTAAGTGGAGGAGTATGGAAAGAGAGTAATTGACAGAGAGAACCCCCACAGGTTACT	2886
Qy	931	GlnThrLeuProGluGlyLysAspAlaHisAspAspMetValThrSerGluValAspPhe	950
Db	2887	GAACTCTGCCAGAGAACAGAGAGGCCCGCGGCGCACGCTCGTTAGTGAGCGGAATTG	2946
Qy	951	ThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeuArgThrGluGluValThr	970
Db	2947	ACCCCGAAGCTGTGCAGCTGCAGAACTGCAGGCGCCATTGGGTGCCGAGAGAGAAC	3006
Qy	971	GluAlaSerGlyAlaGluGluThrThrAspMetValSerAlaValSerGlnLeuThrAsp	990
Db	3007	GAAGATCTGCTGCTGAAGAGACACAGAAATGGTGTGACAGCTCTCCAGGTTAACCGAC	3066
Qy	991	SerProAspThrThrGluGluAlaThrProValGlnGluValGluSerGlyValLeuAsp	1010
Db	3067	TCCCCAGACACACAGAGGAGGCCACTCCGGTGCAGGAGGTGGAAGGTGGCGTACCTGAC	3126
Qy	1011	ThrGluGluGluArgGlnThrGlnAlaIleLeuGlnAlaValAlaAspIysValIys	1030
Db	3127	ATAGAAGACGACAGAGCGCGACTCAAGAGGCTCTCCAGGAGTGTCCAGGAGTGTGCA	3186
Qy	1031	GluGluSerGlnValProAlaThr-----GlnThrValGlnArg	1043
Db	3187	GAGAAATCCACAGCTGCCTGCCACCGGTGGCCAGAGATGCTTCCAGCTGTGCAGAGA	3246
Qy	1044	ThrGlySerLysAlaLeuGluIlysValGluValGluGluAspSerGluValLeuAla	1063
Db	3247	GCA-----GAGGCAGAAAGACCCAGAGCAGGAGTGAAGCGTCGGGT	3288
Qy	1064	SerGluIlysGluIlysAspValMetProLysGlyProValGlnGluAlaGlyAlaGluHis	1083
Db	3289	CTGAAGAAAGACAGCGGATGTAGTGTGTAAGTAGATGCTCAGGAGGCCAAAACTGAGCCT	3348
Qy	1084	LeuAlaGlnGlySerGluThrGlyGlnAlaThrProGluSerLeuGlu---ValProGlu	1102
Db	3349	TTTACAGAGGGAAGGTGTGGGGCAGACCCACCCAGAAAGCTTTGAAAAAGCTCTCAA	3408
Qy	1103	ValThrAlaAspValAspHis-----ValAlaThrCysGlnVal-----	1115
Db	3409	GTCACAGAGCATAGAGTCCAGTGTGATCAACCTTGTCAAGCCGAAACCTTAGCT	3468
Qy	1116	---IleLysLeuGlnLeu---MetGluGlnAlaValAlaProGluSerSerGluThr	1133
Db	3469	GGGTATAAATCACAGGAGATGTGTATGGTAAGCAGGCTATCCCCCTGACTGCTGTGGA	3528
Qy	1134	LeuThrAspSerGluThrAenGlySerThrProLeuAlaAspSerAspThrAlaAspGly	1153
Db	3529	CCTACAGACAGTGTAGACTATGTGAAGCACCCTCCCTAGCGGACTTTGACGCGCACAGG	3588

```

QY 1154 ThrGlnGlnAspGluThrIleAspSerGlnAspSerIysAlaThrAlaAlaValArgGln 1173
D 3589 ACCCGAAGAGAGAGATGTTGAAATTCATGAGAGAGAAATGAGCTGCTGTCACCCAG 3648
QY 1174 SerGlnValThrGlnGlnGlnAlaThrAlaGlnIlySglnGluProSerThrLeuPro 1193
D 3649 TCAGGGGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3705
QY 1194 AsnAsnValProAlaGlnGlnGlnIlySglnGluProGly--ArgAspValLeuGlu 1212
D 3706 TCACATTTTGTGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3765
QY 1213 ProThrGlnGlnGlnLeuThrAlaAlaValProValLeuAlaIlySglnGluValGly 1232
D 3766 CATACAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3825
QY 1233 GlnGlnGlnGlnValAspThrLeuAspGlnIlySglnValIlyS--Gln 1247
D 3826 CAA-----GAGGCTGACAGAGATGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3879
QY 1248 GlnGlnGlnValPheValHisSerGly-----ProAsnSerGlnIlySAlaAspVal 1265
D 3880 GGACTTGGAGGGCTTATAGACACAGCATPAACAGTCAGCGGAGAGAGAGAGAGAGAGAG 3939
QY 1266 ThrTyrAspSerGlnValMetGlyValAlaGlyCysGlnGlnIlySglnGluSerThrGluVal 1285
D 3940 GCCCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3999
QY 1286 GlnSer-----LeuSerLeuGlnGlnGlnIlyMetGlnThrAspValGlu 1300
D 4000 CAGAGTCAAGCTAAGCTCTCTCCATCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4059
QY 1301 LysGlnIlySglnGlnThrIlySglnGlnGlnValSerGlnGlnGly--GlnGlnGlnThr 1319
D 4060 AGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4119
QY 1320 AlaAlaProGlnHisGlnGlnIlySglnGlnIlySglnGlnIlySglnGlnIlySgln 1339
D 4120 GCTGTACCGTATCTGAAAGAGTCAAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4179
QY 1340 SerGlnIlySglnIlySglnIlySglnIlySglnIlySglnIlySglnIlySglnIlySgln 1358
D 4180 ATAGATGGGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4239
QY 1359 AspIlySglnIlySglnIlySglnIlySglnIlySglnIlySglnIlySglnIlySglnIlySgln 1378
D 4240 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4299
QY 1379 AlaGlnAlaValGlnIlySglnIlySglnIlySglnIlySglnIlySglnIlySglnIlySgln 1396
D 4300 GCGGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4359
QY 1397 SerProGlnIlySglnIlySglnIlySglnIlySglnIlySglnIlySglnIlySglnIlySgln 1416
D 4360 AGTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4419
QY 1417 HisTrpThrLeuGlnHisAlaGlnAspThrValProLeuGlnIlySglnGlnIlySglnGlnIlySgln 1436
D 4420 GACTTGGCGCTCATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4479
QY 1437 SerIleProIleIleValThrProAlaProGlnIlySglnIlySglnIlySglnIlySglnIlySgln 1456
D 4480 TCGACACCGATATGATCTCTCTACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4539
QY 1457 GlnIleSerAlaSerGlnIlySglnIlySglnIlySglnIlySglnIlySglnIlySglnIlySglnIlySgln 1476
D 4540 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4599
QY 1477 AspAlaAspGlnIlySglnIlySglnIlySglnIlySglnIlySglnIlySglnIlySglnIlySgln 1496
D 4600 GAGGTC-----AAAGTGAAGTGAAGATTTGAGAGAGAGATTTGAGAGAGAGATTTGAGAGAG 14653
QY 1497 LeuGlnLeuGlnSerIlySglnIlySglnIlySglnIlySglnIlySglnIlySglnIlySglnIlySgln 1516

```

```

D 4654 TTGAACTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4713
QY 1517 GlnPheAlaArgThr--GlnThrAlaProGlnIlySglnIlySglnIlySglnIlySglnIlySgln 1535
D 4714 CAGTTTGAAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4773
QY 1536 ValProAlaCysArgLeuAspSerArgGlnProAsnArgCysTrpThrIlyS-- 1552
D 4774 GCTCAGTGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4833
QY 1553 --MetIlySglnIlySglnIlySglnIlySglnIlySglnIlySglnIlySglnIlySglnIlySgln 1567
D 4834 GAACTCAGGCGCTGCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4893
QY 1568 -----AspLeuGlnValLeuThrValLeuGlnAlaThrAlaGlnIlySglnIlySgln 1583
D 4894 TCACCGCAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4953
QY 1583 eCysLeuProArgLeuGlnIlySglnIlySglnIlySglnIlySglnIlySglnIlySglnIlySgln 1594
D 4954 AAGACCATGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4987

RESULT 7
US-10-133-937-37
; Sequence 37, Application US/10133937
; GENERAL INFORMATION:
; APPLICANT: Khan, Javed
; APPLICANT: Ringner, Markus
; APPLICANT: Peterson, Carsten
; APPLICANT: Meltzer, Paul
; TITLE OF INVENTION: METHODS FOR ANALYZING HIGH DIMENSIONAL DATA FOR CLASSIFYING,
; TITLE OF INVENTION: DIAGNOSING, PROGNOSTICATING, AND/OR PREDICTING DISEASES AND
; FILE REFERENCE: 11613.56US01
; CURRENT APPLICATION NUMBER: US/10/133,937
; CURRENT FILING DATE: 2002-11-04
; NUMBER OF SEQ ID NOS: 99
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 37
; LENGTH: 9161
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-133-937-37

Alignment Scores:
Pred. No.: 1,41e-12 Length: 9161
Score: 492.00 Matches: 386
Percent Similarity: 33.47% Conservative: 269
Best Local Similarity: 19.72% Mismatches: 714
Query Match: 6.09% Indels: 589
Gaps: 75

US-09-902-432-4 (1-1596) x US-10-133-937-37 (1-9161)
QY 45 AlaAspProAlaThrIlySgln-----ProGlnIlySglnIlySglnIlySglnIlySglnIlySgln 59
D 1399 GCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1458
QY 60 SerValAsnGlnIlySglnIlySglnIlySglnIlySglnIlySglnIlySglnIlySglnIlySgln 73
D 1459 ATCTCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1518
QY 74 GlnIlySglnGlnIlySglnIlySglnIlySglnIlySglnIlySglnIlySglnIlySglnIlySgln 93
D 1519 CAAAGAGATTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1578
QY 94 SerGlnAspValArgGlnIlySglnIlySglnIlySglnIlySglnIlySglnIlySglnIlySglnIlySgln 113
D 1579 AGGCTGATAGCGGAGAGAGT-----CTGAAGCAGAGCGGAGAGAGAGAGAGAGAGAGAGAG 1629
QY 114 GlnAspIleThrIlySglnIlySglnIlySglnIlySglnIlySglnIlySglnIlySglnIlySglnIlySgln 133

```

Db 1630 AATCCGTGCGCAAGGAGTCAAAAGAAACCCCTGAGGTC-----ACA 1674
 Qy 134 SerGluAsnValGluGluMetValGlnProAlaGluSerGlnAlaAsnAspValGly 153
 Db 1675 AAAGTGAATCAGCTGGAA-----AAGCCACCAAAAGTTGAAGCAAAGAA----- 1719
 Qy 154 PheLysLysValPheLysPheValGlyPheLysPheThrValLysLysAspLysAsnGlu 173
 Db 1720 -----AAGTAATGGTGAAGAAAGCAACGCCAGTA 1749
 Qy 174 LysSerAspThrValGlnLeuLeuThrValLysLysAspGluGlyGluGlyAlaGluAla 193
 Db 1750 AAAACAGAGACCAAACTTCAGTGACTGAAAGGAG----- 1785
 Qy 194 SerValGlyAlaGlyAspHisGlnProSer---ValGluThrAlaValGlyGluSer 212
 Db 1786 -----GTTCCCAAGCAAGAGGACCATCTCCAGTGAAGCGGAGTGGCTGAG--- 1833
 Qy 213 AlaSerLysGluSerGluLeuLysGlnSerThrclLysGlnGluGlyThrLeuLysGln 232
 Db 1834 -----AACAAGCCACAGATGTCAAACCCAAAGCTGCCAAGGAG 1872
 Qy 233 GluGlnSerSerThrGluIleProLeuGlnAlaGluSerAspGlnAlaAlaGluGluGlu 252
 Db 1873 AAGCGGTGAAAAAGGAAACAAAGGTAAAGCTGAACAGAAAGAGGAGAAAGAAAG 1932
 Qy 253 AlaLysAspGluGlyGluGlnGlnLysGlnLysGluProThrLysSerProGluSerPro 272
 Db 1933 CCAAGAAAGAAAGTGGCTAAAGAGGAGCAAAACACCTATCAAGAGAGGAGGAAACCA 1992
 Qy 273 SerSerProValAsnSerGluThrThrSerSerPheLysLysPheThrHisGlyTrp 292
 Db 1993 AAA-----AAGGAAGAGGTGAAAAAGAAAGTCAAAAAGAGATC----- 2031
 Qy 293 AlaGlyTrpArgLysLysThrSerPheLysLysSerLysGluAspAspLeuGluThrAla 312
 Db 2032 -----AAGAAGAGAGAGAAAAAGAACCCCAAGAGAGGTTAAGAAAGAAACCCG 2082
 Qy 313 GluLysArgLysGluGlnGluAlaGluLysValAspGluGluLysGluLysThrGlu 332
 Db 2083 CCAAGGAAGTCAAGAGGAAGTTAAGAG-----GAAGAGAAAGAGAA 2127
 Qy 333 ProAlaSerGluGlnGlnProAlaGluAspThrAspGlnAlaArgLeuSerAlaAsp 352
 Db 2128 GTGAAAAAGAGAAAGAAAGCAACCCAAAAAGAAATT-----AAGAAGCTCCCTAAAGAC 2181
 Qy 353 TyrGluLysValGluLeuProLeuGluAspGlnValGlyAspLeuGluAlaSerSerGlu 372
 Db 2182 GCAAGAAATCATCTACTCTCTG-----TCTGAAGCA 2214
 Qy 373 GluLysCysAlaProLeuAlaThrGluValPheAspGluLysMetGluAlaHisGlnGlu 392
 Db 2215 AAAAAACCACTGCTTTAAACCAAAAGTACCCCAAGAGGAGAGTCTGTCAAGAAAGAT 2274
 Qy 393 ValValAlaGluValHisValSerThrValGluLysThrGluGluGlnGlyGly 412
 Db 2275 TCTGTGCTCCGGAAGCCAAAGGAGAGAGGGGAAAAATAAAAGTCATTAAAGAGGAAGGC 2334
 Qy 413 GlyGluAlaGluGlyGlyValValValGluGlyThrclLysLeuLeuProGluLys 432
 Db 2335 AAGGCCGAGAGGCTGCGCTGAGCTGTGCGCACTGAGCCACCAAGAGCTGTGATG 2394
 Qy 433 LeuAlaGluProGlnGluValProGlnGluAlaGluProAlaGluLeuMetLysSer 452
 Db 2395 GCGGCACTGGAATAGAGCC-----ATTGGCCCTGCCAAAGAACTCGAAGCTGAG 2445
 Qy 453 ArgGluMetCysValSerGlyGlyAspHisThrGln---LeuThrAspLeuSerProGlu 471
 Db 2446 AGGTCCCTTATGTCATCTCTCTGAGGATCTAACCAAGGACTTTGAAGAGTTAAAGGCTGAA 2505
 Qy 472 GluLysThrLeuProLysHisProGluGlyIleValSerGluValGluMetLeuSerSer 491
 Db 2506 GAGGTGATGTAAACAAG-----GACATCAAGCCTCAGCTGGAGCTTAATCGAAGAC 2556

Qy 492 GlnGluArgIleLysValGlnGlySerProLeuLysLysLeuPheSerSerSerGlyLeu 511
 Db 2557 GAAGAGAAACTGAAG---GAAACTGAGCCAGTCCGAAGCTACGTCTATCCAGAAGGAGAGA 2613
 Qy 512 LysLysLeuSerGlyLysLysGlnLysGlyLysArgGly----- 524
 Db 2614 GAAGTCAACCAAGGTCTCGCGAGTCCCCTGATGAGGGAATCACTACCACTGAAGGGGAG 2673
 Qy 525 GlyGlyLysAspGluProGlyGluTyrGlnHisIleThrGluSerProGluSer 544
 Db 2674 GCGCAATGTGAACAGACACTCTGAGGAGTGGAGCCCGTCGAGAGCAGGAGGTAGAC--- 2730
 Qy 545 AlaAspGlnLysGlyLysSerSerAlaSerProGluProGluProGluThrThr 564
 Db 2731 ---GACATTGAAAAATTTGAAGATGAAGGAGCGCGTTTGAAGAATCTTCAGAGACT--- 2784
 Qy 565 CysLeuGluLysGlyProLeuGluAlaProGlnAspGlyGluAlaGluGluGlyThrThr 584
 Db 2785 -----GGAGACTATGAAGAGAGGAGCAACTGAGGAGGCTGAGGAGCCAGAA 2832
 Qy 585 SerAspGlyLysLysArgGluGlyIleThrProTrpAlaSerPheLysLysMetVal 604
 Db 2833 GAGGATGGGAGGAACAC-----GTA 2853
 Qy 605 ThrProLysLysArgValArgArgProSerGluSerAspLysGluGluLeuGlu--- 623
 Db 2854 TGTGTGAGCGCTCCAAGCACACGCCCACTGAGGATGAGGAAAGTCCCAAGCGGAGGCT 2913
 Qy 624 -----LysValLysSerAlaThrLeuSerSerThrAspSerThrValSerGlu 639
 Db 2914 GATGCATACATCAGGAGAGAGGGAGTCTGTGCCCACTCTGAGGAGGCTGACCGAGCGAAGAA 2973
 Qy 640 MetGlnAspGlnValLysThrValGlyGluGlnLysProGluProLysArgArg 659
 Db 2974 GACATGATGAGGCCATTGAGAAGGAGAGGCTGAACAACTCTGAAGAGGAGGCTGATGAG 3033
 Qy 660 ValAspThrSer-----ValSerTrpGluAlaLeu 669
 Db 3034 GAGGACAAAGCTCAAGATGCCAGAGAGGAGGAATATGAGCGCGAAAAAATGGAAGCTGAA 3093
 Qy 670 IleCysValGlySerSerLysLysArgAlaArgLysAlaSerSerSerSerAspGluGly 689
 Db 3094 GACTATGTGATGCTGTGTCGCAAGGCTGCAAGGCTGTGTGTCGCGAGGAGCAGTAT 3153
 Qy 690 Gly-----ProArgThrLeuGlyGlyAspSerHisArgAlaGluGluAla 704
 Db 3154 GGATTCCTCACCAACCAAGCAACTAGGAGCCCACTCTCTGGCCGAGAACCTTGCA 3213
 Qy 705 SerLysAspLysGluAlaGlyThrAspAlaValProAlaSerThrGlnGluGlnAspGln 724
 Db 3214 TCTTCAATTCATGAT-----GAGACTTTACCTGGAGGCTCAGAGCGAGGCCACC 3264
 Qy 725 AlaGlnGlySerSerSerProGluProAlaGlySerProSerGluGlyGluGlyValSer 744
 Db 3265 GCTTCTGATGAGGAGAAATCGAGAA-----GACCAGCCTGAGGAATTCACCTGCACCTCT 3318
 Qy 745 ThrTrpGluSerPheLysArgLeuValThrProArgLysLysSerLysSerLysLeuGlu 764
 Db 3319 GGCTAC-----ACTCAGTCTACTATTGAG 3342
 Qy 765 GluLysAlaGluAspSerSerValGluGlnLeuSerThr----- 777
 Db 3343 ATATCCAGTGAGCCCAACCCCATGGATGAGATGTCTACCCCTCGAGACGTGATGATGAT 3402
 Qy 778 -----GluIleGluProSerArgGluSerTrpValSerIleLysLys--- 792
 Db 3403 GAGACCAACAATGAAGAGAGCGGAGTCCCTTCTCAGGAATTCGTAATAATATCACCAATAT 3462
 Qy 792 ----- 792
 Db 3463 GAATCTTATTGTTTCTCAGGAATACTCTAAACCTGCTGATGTTTACACCGCTCAACGGA 3522

QY 793 PheileProGlyArgArgLysArgAlaAspGlyLysGlnGlnAlaThrValGlu 812
 Db 3523 TTTTCTGAAGATCAAAAACAGATGCCACTGATGGCAAGATTACAAATGCTTCAGCTCT 3582
 QY AspSerGlyProValGluIleAsnGluAspAspProAsnValProAlaValProLeu 832
 Db 3583 ACCAATATCACCAACC----- 3597
 QY 833 SerGluThrAsnAlaValGluArgLysMetGluAlaGlnGlyAsnThrGluLeuPro 852
 Db 3598 -----TCTTCATGAGAGAACAAATTCAGACATGCT----- 3633
 QY 853 GlnLeuLeuGlyAlaValThrValSerGluGluLeuSerLysThrLeuValHisThrVal 872
 Db 3634 -----TTACGTGATGCTTACTGCTGTGAAGTGAAGAACGACGACC----- 3672
 QY 873 SerValAlaValIleAspGlyThrArgAlaValThrSerValGlu-----GluArg 889
 Db 3673 ACTTGGACATCAAGATAGCATCTCAGCTGTTTCAAGTGAAGAAAGTCAAGCCATCGAAG 3732
 QY 890 SerProSerTrpIleSerAlaSerValThrGluProLeuGluHisThr---AlaGlyGlu 908
 Db 3733 AGCCCGTCC---CTGAGTCATCTCCACATCATCACCTTGAAGAACCCCGGCTGA 3789
 QY 909 -----AlaMetProProValGlu---GluValThrGluLysAspIleIle 922
 Db 3790 CCTAGTGTCACCTCTCTCTGACGCCCAATGAGATTAAAGTCTCTGACAGACAGAAATGA 3849
 QY 923 AlaGluGluThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHisAspAsp 942
 Db 3850 GCGCCGGTCTCTCTGAGGTGACCAAGATGTTGAAGAACATTTGCTAGTCTGCTGAG 3909
 QY 943 MetValThrSerGluValAspPheThrSerGluAlaValThrAlaThr----- 958
 Db 3910 GACAAGACTCGAAGGTGTGCACCATCTCAGTCCGTGACGTGGCGTCTGTCACACA 3969
 QY 959 -----GluThrSerGluAlaLeuArgThrGluGluVal--- 969
 Db 3970 CCTTACTATCAATCTCTCTGACGAGAAATCCAGTATCTCTTACAGAAATCATTTGAA 4029
 QY 969 ----- 969
 Db 4030 AAACACACGACGTTCCAGTGAGTTTGAATTCACTGATGCCAAAGATGAATGAAG 4089
 QY 969 ----- 969
 Db 4090 GCTTCAGTAAGCCCATGATGAGCCGTCGATCAGAGTCTCTATTGAAAAAGTT 4149
 QY 969 ----- 969
 Db 4150 TTGTCTCTTACGAGCCGCCCTCATTTGATCCGAGTCTGCTTATGAAATTTCTTA 4209
 QY 970 -----ThrGluLysSerGly 974
 Db 4210 ACTGCTGATGACAGGCTTCTGCGCAGAGTCCGAAAGTCTTTTGAAGAAAAAGTGA 4269
 QY 975 AlaGluGluThrThrAspMetValSerAlaValSerGlnLeuThrAspSer----- 991
 Db 4270 AAACAGAGCTCTCCAGACCAAGTAATCTCAAGTTTCTGAATAGCTTCTCTAGTCTTAC 4329
 QY 992 ProAspThrThrGlu-----GluAlaThrProValGlnGluValGluSerGly 1007
 Db 4330 CAAGACAAACAGAGAAAGCAAGACAGACTTTCACCATTAATAAAGAACACTTTGGCCAA 4389
 QY 1008 ValLeuAspThrGlnGluGluGluArgGlnThrGlnAlaIleLeuGlnAlaValAlaAsp 1027
 Db 4390 GAAAAAGAAATGATGATGTGAAGCCATGAGTTTCAACACGACACTGGCTCGATGA 4449
 QY 1028 Lys----- 1028
 Db 4450 AGGAATTTAGAGATTTCTCCACACAAATGATGTCAGTCAAGTTTGATCTTTAA 4509
 QY 1029 -----ValLysGluGluSerGlnValProAla 1037

Db 4510 GAAGACTAAGATGTCATTTCTGAAAGTACTGTCTTCAGCAAGTCCACTCTGTT 4569
 QY 1038 ThrGlnThrValGlnArgThrGlySerLysAlaLeuGluLysValGluGluValGluGlu 1057
 Db 4570 GATGAGGCGGTAGACAGAAAGACACGTACTCTCATATGAGGGGTGGCTTCACCA 4629
 QY 1058 AspSerGluValLeuAlaSerGluLysGlu-----LysAspValMetProLysGly 1074
 Db 4630 GCCTCAGGGCTACGAGCTCATTTCCAGACCAACACAGATGATGTCTCCATCTCTG 4689
 QY 1075 ProValGlnGluAlaGlyAlaGluHis----- 1083
 Db 4690 CATGCT---GAGGTGGCTCCACATTCACAGAAATGATGACTCCTTCAGTGTCT 4746
 QY 1084 LeuAlaGlnGlySerGluThrGlyGlnAlaThr----- 1094
 Db 4747 GTTGGCAACACCTACACCAATTCAGGAACAGAAATGTCATTAAGAAATGC 4806
 QY 1095 -----ProGluSerLeuGluValProGluVal-----Thr 1104
 Db 4807 CCAGACCGATGTCATTTCTCCACAGATTTCTCCCTAAACTGCAAAAGTCCAGACACA 4866
 QY 1105 AlaAspValAspHisValAlaThrCysGlnValIleLysLeu-----GlnGlnLeu 1121
 Db 4867 CCGTTCAAGATCACAGATCTGAACAGTCTCTCAATGCTTATTAATTTGGCCAGAAATCT 4926
 QY 1122 MetGlnAlaValAlaProGluSerSerGluThrLeuThrAspSerGluThr----- 1139
 Db 4927 CCGAGCAATCCCTTGATGAGCTTCAGTCCAGACGTCCAGATCACCCACAGTGGT 4986
 QY 1140 -----AsnGlySerThrProLeu-----AlaAspSerAsp 1149
 Db 4987 GCAGCGCTCTCATCATCTGAATGAGCAATGAGCACTGACAGTCTCTTAC 5046
 QY 1150 ThrAlaAspGlyThr-----GlnGlnAspGluThrIleAspSerGln 1163
 Db 5047 ATGACGACCTCCAGTTTATTCACATTAATACCACTATGAGAGCCCTCTACCCCA 5106
 QY 1164 AspSerLysAlaThrAlaAlaValArgGlnSerGlnValThrGluGluGluAlaAlaThr 1183
 Db 5107 GATATGATCTTTCTGACATCTCAGTATCTCAAGTA---GAGGCTCCCGTCCAC 5163
 QY 1184 AlaGlnLysGluGluProSerThrLeuProAsnAsnValProAlaGlnGluGluHisGly 1203
 Db 5164 TCTTCTGCTCATACCTCTCTCAGATCCCTCT-----CCTCTCCA----- 5205
 QY 1204 GlnGluProGlyArgAspValLeuGluProThrGlnGlnGluLeuThrAlaAlaVal 1223
 Db 5206 GAAGATCTTATCCGATGTGCTCTCCAGATATATGCTTATATGCTTCACTCAC 5265
 QY 1224 ProValLeuAlaLysThrGluValGlyGlnGluGlyValAspTrpLeuAspGlyGlu 1243
 Db 5266 -----TCTGAAAAGTCCAAAGTCTGGAAAGAGAG 5295
 QY 1244 LysValLysGluGlnGluValPheValHisSerGlyProAsnSerGlnLysAlaAla 1263
 Db 5296 AAGCTCTCCAAATCTGAATATCTCTCACCTCACCCACAGAGTCTCTCTTATAT 5355
 QY 1264 AspValThrThrAspSerGluValMetGlyVal-----AlaGlyCysGlnGlu 1279
 Db 5356 TCACCTACTTTTTCAGATTCTACCTCTGACGACAAAGAAAGCAACTTGGCCACAGT 5415
 QY 1280 LysGluSerThrGluValGlnSerLeuSerLeuGluGluGluMetGluThrAspVal 1299
 Db 5416 TCCTCTTCCACCAATAGATGACGATCCGAGAGCCCTATGCTTCCGCTTCAGTG 5475
 QY 1299 ----- 1299
 Db 5476 TTATTCATACATGACCAACACATCTAGCTTGAATAGATTGTCCACACTGGCTG 5535
 QY 1300 GlnLysGluLysArgGluThrLysProGluGlnValSer-----GluGlnGluGlu 1316

Db 5536 GAGAAGGACAGTGGAGGAAGACACACCTGGTGACTTTAGCTATGCCTATCAAAAGCCTGAG 5595
 Qy 1317 GlnGluThrAlaalaProGluHisGluGly-----ThrTyGlyLysProVal 1332
 Db 5596 GAAACACACAGGTGCTCCAGATGAGAGAGATTATGACTATGAGTCTTATGAGAACACACC 5655
 Qy 1333 LeuThrLeuAspMet-----ProSer 1339
 Db 5656 CGACCTCAGATGGGTGGCTATTACTATGAGAAGATGAGAGAACCAAAATCTCCA 5715
 Qy 1340 SerGluArgGlyLysAlaLeuGlySerLeuGlySerProSerLeuProAspGlnAsp 1359
 Db 5716 AGTGACAGTGGCTACTCTATGAGACCATGGGAAACTACCAAGACCCCTGAGATGGT 5775
 Qy 1360 LysAlaGlyCysIleGluValGlnValGlnSerLeuAspThrThrValThrGlnThrAla 1379
 Db 5776 -----GACTATTCCTAT 5787
 Qy 1380 GluAlaValGluLysValIleGluThrVal-----ValIle 1391
 Db 5788 GAAATTATTGAGAAGACCAACCGGACCCCTGAAGAGGTGGGTACTCATATGACATAAGT 5847
 Qy 1392 SerGluThrGlyGluSerProGluCysValGlyAlaHisLeuLeuProAlaGluLysSer 1411
 Db 5848 GAAAAGACCAACCGCCCGGAGTGGTTACAGCTATGAAAGACTGAGAGGTCT 5907
 Qy 1412 -----SerAlaThrGlyHisTrp 1418
 Db 5908 AGAAGGCTTCTGGATGACATCAGCAATGGCTATGATGACTCTGAGTGGTGGCCAC--- 5964
 Qy 1419 ThrLeu-----GlnHisAlaGluAspThrValProLeuGlyProGlu 1432
 Db 5965 ACACCTGGGAGCCCGAGCTACTCTTATGAAACCACTGAGAAATACCAGTTTCCCTGAG 6024
 Qy 1433 SerGlnAlaGluSerIleProIle----- 1441
 Db 6025 TCTGAGGTATTCTCTATGAGACATCTACAAGACACACGACCCCTGTACTTCCACA 6084
 Qy 1442 -----ValThrProAlaProGluSerThrLeuHisProAsp 1453
 Db 6085 TACTGTTACGAGCTGCAGAGAAAATCACTAGAACCCCTCAGGCATCCACATATTCCTAC 6144
 Qy 1454 LeuGlnGlyLysSerAlaSerGlnArgGluArgSerGluGluAspLysProAsp 1473
 Db 6145 GAGACTTCAGACCTATCTACTACTCGAGAAAAGAGTCCCTCCAGAGCCGCTCAGGAT 6204
 Qy 1474 AlaGly-----ProAspAlaAspGlyLysGlu 1482
 Db 6205 GTCGATTATGCTCGTGCTCTTGTGAATACAGCACCCCAAGACAGAGCTTTCACCC 6264
 Qy 1483 Ser-ThrAlaIleGluLysValLeu----- 1490
 Db 6265 TCATTATTATCCCAATCCTCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 6324
 Qy 1491 -----LysAlaGluProGluIleLeuGluLeuGluSerLysSerAs 1504
 Db 6325 GAAAAGCCCTCACTCAATCAGGGGAGCCCGCTCCAGGAGGAAACCAACAGGC 6384
 Qy 1504 nLysIleValLeuAsnValIleGlnThrAlaValAspGlnPheAlaArgThrGluThr-- 1523
 Db 6385 CGACAGTGTGATGAACCCCTCC-ACCTCAGTCAGCGAGTCAGCCCTCCAGACCCGA 6443
 Qy 1524 -----AlaProGluThrHisAlaTyAspSerGlnThrGlnValProAlaCysAr 1540
 Db 6444 CTCTGATGTTCCCGGAGATGAAGTGCCCTCCATCAG-----GCCGATGCCAA 6497
 Qy 1540 gLeuAspSerArgGluProAsnArgCysTrpThrLysMetLysAspAlaLysMetLysHi 1560
 Db 6498 TATCGACTCTGAAGACAGTGCAGAAACCACTCCCAACAGACAACTGTCTACGTACAAACA 6557
 Qy 1560 s-----ProVal-----ProGlnProArgGlu----- 1567
 Db 6558 CATGGACCCACCTCCAGCTCCCGTGAAGACCGCAGCCCTTCGCCACCGCCCTGATGT 6617

Qy 1568 -----AspLeuGlnValLeuThrValLeuGlu-AlaTrpAlaGln----- 1580
 Db 6618 GTCCATGTGGACCCAGAGCCCTTGCCTATGACGAGAACCTGGGCAAGCTCTAAAGAA 6677
 Qy 1581 -----ProArgLysCysLeuProArgLeuGlnLeuLysAlaPro 1593
 Db 6678 AGATCTGAAAGAGAGACCAACCAAAAGCCAGGTACAAAGACCA 6724

RESULT 8
 US-10-203-138A-6112
 ; Sequence 6112, Application US/10203138A
 ; GENERAL INFORMATION:
 ; APPLICANT: Molecular Dynamics, Inc.
 ; APPLICANT: Penn, Sharron G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Wenheng
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 ; FILE REFERENCE: PB 0004 WO 8
 ; CURRENT APPLICATION NUMBER: US/10/203,138A
 ; CURRENT FILING DATE: 2002-08-02
 ; PRIOR APPLICATION NUMBER: US 60/180,312
 ; PRIOR FILING DATE: 04 February 2000 (04.02.00)
 ; PRIOR APPLICATION NUMBER: US 60/207,456
 ; PRIOR FILING DATE: 26 May 2000 (26.05.00)
 ; PRIOR APPLICATION NUMBER: US 09/632,366
 ; PRIOR FILING DATE: 03 August 2000 (03.08.00)
 ; PRIOR APPLICATION NUMBER: GB 24263.6
 ; PRIOR FILING DATE: 03 October 2000 (03.10.00)
 ; PRIOR APPLICATION NUMBER: US 60/236,359
 ; PRIOR FILING DATE: 27 September 2000 (27.09.00)
 ; PRIOR APPLICATION NUMBER: US 60/234,687
 ; PRIOR FILING DATE: 21 September 2000 (21.09.00)
 ; PRIOR APPLICATION NUMBER: US 09/608,408
 ; PRIOR FILING DATE: 30 June 2000 (30.06.00)
 ; NUMBER OF SEQ ID NOS: 15438
 ; SOFTWARE: Molecular Dynamics Sequence Listing Engine
 ; SEQ ID NO 6112
 ; LENGTH: 7997
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: MAP TO AL034555.2
 ; FEATURE:
 ; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 12
 ; FEATURE:
 ; OTHER INFORMATION: EST_HUMAN HIT: AU117052.1, EVALUE 0.00e+00
 ; FEATURE:
 ; OTHER INFORMATION: SWISSPROT HIT: P08640, EVALUE 4.00e-10
 ; FEATURE:
 ; OTHER INFORMATION: NT HIT: g11427711, EVALUE 0.00e+00
 ; US-10-203-138A-6112

Alignment Scores:
 Pred. No.: 4,03e-10 Length: 7997
 Score: 431.50 Matches: 373
 Percent Similarity: 34.27% Conservative: 271
 Best Local Similarity: 19.85% Mismatches: 759
 Query Match: 5.34% Indels: 476
 DB: 6 Gaps: 83

US-09-902-432-4 (1-1596) x US-10-203-138A-6112 (1-7997)

Qy 31 ProAlaAlaGluAlaSerGlyAlaAlaGly-AspProAlaAspAlaAspProAla----- 48
 Db 886 CTGGACCAAGAACTCAGGTCTCTCAGACCGAGCGCTGCAAAATCTGACTTGTCTAAACT 945

Qy 49 -----ThrLysLeuProGlnLysAsnGlyGlnLeuSerValAsnGlyVa 64
 Db 946 GGAATCAGTTAGATGAAGTAGTACCAAGAGAAAGGGGCTTTCAAGCCATGTTGAAGTGT 1005

Db 3025 GAATAAAGATTTCAGAACTGAAACTCCACCTTCCTGGGCGCTCCAGTGTCTACAGTCGT 3084
 Qy 647 lGlyGluGluGluProGluGluProLysArgArgValAspThrSerValSerTrpGl 667
 Db 3085 AACTCTAGATACAGCCCATCAGCACTAGAGAAAGACCACTGGTGCAAAACCGGTAGAGGC 3144
 Qy 667 uAlaLeuLeuLeuLeuSerSerSerSerSerSerSerSerSerSerSerSerSerSer 687
 Db 3145 GCCTTTGGTA-----ACAGAAGGAAGACTGTGGAGCCAGCTACCGCTCTCAGAAGA 3195
 Qy 687 p-----GluGlyGlyProArgThrLeuGlyGlyAspSerHisArgAlaGl 702
 Db 3196 AGCAAGCCTGCATCTGAACCTCTCTGCCCCCTGTG-----GAACAGCTGGA 3243
 Qy 702 uGlu-----AlaSerLysAspLysLysGluAlaGlyThrAspAlaValPr 716
 Db 3244 ACAAGTAGACCTGCCCCCAGGAGCAGACCCCGATAAAGAGCTGCC-----ATGATGCC 3297
 Qy 716 oAlaSerThrGlnGluGlnAspGlnAlaGlnGlySerSerSerProGluPro----- 733
 Db 3298 TGGCGGTGTGAGAA-----GGTTATCATCAGGTGACCGCGCGCTTATCT 3342
 Qy 734 ----AlaGlySerProSerGluGlyGluGlyValSerThrTrpGlu----- 747
 Db 3343 GGATGCCAAGCTTCAACTCCCGGGGCTCGTTTCCCGGAGAGCAACGTAGATCC 3402
 Qy 748 -----SerPheLysArgLeuValThrProArgLysLysSerLysLysLeuGl 764
 Db 3403 AGAGCCTGACAGTACCAGCAGCACTTTCAAAACCCAGCTCAGAGTCTGAGGAAGCAATGA 3462
 Qy 764 uGluLysAlaGlu-----AspSerSerValGlu----- 773
 Db 3463 GCCAAAGGCGGAAAGCCAGACCCACTGCAGATGCTGAGCCTGATGCAAAACAGAAAGC 3522
 Qy 774 -GlnLeuSerThrGluLeuGluProSerArgGluGluSerTrpValSerLysLysLysPh 793
 Db 3523 CGAAGCTGCTCTGAGTCTCAGCCCCCAGCTTCTGAAGAT---TTAGAGTTGATCTCTCC 3579
 Qy 793 eileProGlyArgArgLysLysArgAlaAspGlyLysGlnGluGlnAlaThrValGluAs 813
 Db 3580 AGTTGTGCAAGAGATAAAAGCAAAACAAAGCAAGCGTTCAAGAGCCCTGTTTCAG-- 3637
 Qy 813 pSerGlyProValGluLeuAsnGluAspProAsnValProAlaValProLeuSe 833
 Db 3638 -GCAGTGCAGTGAATCTGTGAG-----AAGCCCGTCAAGGAAGAG 3681
 Qy 833 rGluTyrAsnAlaValGluArgGluLysMetGluAlaGlnGlnLysThr-----GluLe 851
 Db 3682 TGAG-----AGATAGACCGGAAATACTCAAGCGGTCCAATTCCTCGGGGAGAGC 3735
 Qy 851 uProGlnLeuLeuGlyAlaValTyrValSerGluGluLeuSerLysThr----- 867
 Db 3736 ACAGAAGCTTTTGGAAATTGAAGATGGAGCAGAGAGAGATTACAAGGACTGCTTCTAATA 3795
 Qy 868 -----LeuValHisThrValSerValAlaValAlaLysGlyThrArgAla-- 882
 Db 3796 CTCTGCTGCAGACCTTGAAATCCCGAACCAAGTTTGCCTCTCAGCGCAACAGGCGCCG 3855
 Qy 883 ----ValThrSerVal-----GluGluArgSerProSerTrpIl 894
 Db 3856 GAATGTAGAGGCGCTTATCCCAACCATGGGTGACCATGAAACCGCTCTCCT----- 3907
 Qy 894 eSerAlaSerValThrGluProLeuGluHisThrAlaGlyGluAlaMetProValGl 914
 Db 3908 -----GTCAAAGAGCCCGTTGAGCAA-----CC 3930
 Qy 914 uGluValThrGluLysAspIleAlaGlu-----GluThrProValLeuThrGlnTh 932
 Db 3931 AAGAGTGACCAAGAGAGATTGGAGCGAGAGCTTCAGAGGCTGTCAGCGGTTCCCAACCAC 3990
 Qy 932 r-----LeuProGluGlyLysAspAlaHisAspAspMetValThrSerGl 947
 Db 3991 CCTCGAGGGGAGGCGCTTCCAAAGACACGCGCGGAGCGCATGAAGAGGAGGAGAACGA 4050

Qy 947 uValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeuArgThrGl 967
 Db 4051 GGCCAAAGGAACCTGCAGAAACACACTCAAGCCACTGAGGGATGCGCGTCCCAAGGTGCCA 4110
 Qy 967 uGluValThrGluAlaSerGlyAlaGlu-----GluThrThrAspMetValSe 983
 Db 4111 GAAACTGCAGCTGGTGGACCCCAAGGGAAGAAAGGAAAGAAATGAACCGAGGTGGA 4170
 Qy 983 rAlaValSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGl 1003
 Db 4171 TGCTACA-----CGTCTGAGGCCACCACTGAGGTGGGCCCCCAATAGG 4215
 Qy 1003 uValGluSerGlyValLeuAspThrGluGluGluArgGlnThrGlnAlaLysLeuGl 1023
 Db 4216 CBTGAAGAGAGAGCTCCATGGAACCCCAAGGCTGCTGAGGAGGAGGAGGAGTGAACAGAA 4275
 Qy 1023 nAlaValAlaAspLysValLysGluGluSerGlnValProAlaThrGlnThrValGlnAr 1043
 Db 4276 ACCTGACAAAGATGCTGGCAGACAGAAACCCCTGAAACCCGCCCTGTTGAA-- 4333
 Qy 1043 gThrGlySerLysAlaLeuGluLysValGluGluValGluGluAspSer----- 1059
 Db 4334 -----GTGTAGAGAAACCCGCGCCCTGAAACAACTCCAAATCAAAGAG 4380
 Qy 1060 -----GluValLeuAlaSerGluLysGluLysAs 1069
 Db 4381 AGGAAGATCTCGAAACTCCAGGTTAGCAGTGCACAAATCTGCAAGTCTGAAATGTGGA 4440
 Qy 1069 p-----ValMetProLysGlyProValGlnGluAlaGlyAlaGluHisLeuAlaGlnGl 1087
 Db 4441 TGTGCTGTGCTGAGTCCAGGGGGCTGCAGCAGGAGGGGAGAG-- 4486
 Qy 1087 ySerGluThrGly-----GlnAlaThrProGluSerLeuGluValProGluValThrAl 1105
 Db 4487 ----GNAATCTGGGTGCTGGCAGTCTCCCTCTGAGAAAGTGAGAGTCCCAAAAGAGGA 4542
 Qy 1105 aAspValAspHisValAlaThrCysGlnValLysLeuGlnGlnLeuMetGluGlnAl 1125
 Db 4543 TGTGTTTATCA-----TCCAGTGTGAAAGTGATCC 4572
 Qy 1125 aValAlaProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrProLe 1145
 Db 4573 AGTTGATCAGACCAAGAACCCAGAGAAAGACAGCGTGTCTGCTCTGCGGCCGTCGCCAGA 4632
 Qy 1145 uAlaAspSerAspThrAlaAspGlyThrGlnGlnAspGluThrIleAspSer----- 1162
 Db 4633 AGCC---ACCCAGTTAGCCAGCAGATGGAGCTGGAGAGGCGGTGGAACACATCCCAAA 4689
 Qy 1163 ----GlnAspSerLysAlaThrAlaAlaValArgGlnSer----- 1174
 Db 4690 GCTCGCTGAGGCTCTGCTCTGCTTATAAGGCAGATGCACAGAGGGCGCTTGCCTCC 4749
 Qy 1175 -----GlnValThrGluGluGluAlaAlaThrAlaGlnLys 1186
 Db 4750 AGAGGACAGGACCAAGCCCTGCACACCAAGCAAGTGAACAGAGCTGCTGCGGCCATCGG 4809
 Qy 1186 s-----GluGluProSerThrLeuProAsnValProAlaGlnGluHis-- 1202
 Db 4810 CTCATCATCAATGACATTTCTGGGAGCCAGAAACCTTCCAGCAGACCTCCCATATTC 4869
 Qy 1203 -GlyGluGlu-----ProGlyArgAspValLeuGluProThrGl 1215
 Db 4870 TGGAGAAATCCAGACAGATCTCAACCCCGCAGGTGTCAGCGGCTGCAGGCTTCTGA 4929
 Qy 1215 nGlnGluLeu-----ThrAlaAlaAlaValProValLeuAlaLysThrGluValGlyGlnGl 1234
 Db 4930 GGAAGGAATGGAGACAGATGAGGCTGTATCTGGCATCTCTGGAACCTGAGGCTGCTACAGA 4989
 Qy 1234 uGlyGluValAspTrpLeuAspGlyLysValLysGluGluGlnGluValPheValHi 1254
 Db 4990 ATCTTCTAGGCGCTCCAGTCAATGCTGACCCCTCA----- 5026

```

QY 1254 ssergLyProaensergLInLysAlaAlaAspValThrTyraAspsergLuValMecLyVal 1274
Db 5027 -GCCGCCCCACAGATACCAAGAGAGCCAGAGAAATACAGTGAACCTCACACTCAGT 5085
QY 1274 lAlaGlyCysgIngluLysgIserThrgLValAlGInserLeuSerLeuGluGluGly 1294
Db 5086 GCCAGAGAGCCAAAGGCTCAAGAAAGTGAAGTCACTCTGTTCCGAAAGACAAAGGGGG 5145
QY 1294 uMergLuthraAspValGluLysgIuLysAlrgLuthrLysProgluGluValserGluGly 1314
Db 5146 CCAGAAACACCCCGATCAAGCGCAAGGAAACACAAACAGAAAGAGTGGTGGCT----- 5200
QY 1314 uGlyGluGluGluThrAlaAlaProgluHIsGluGlyThrTyraGlyLysProValLeuth 1334
Db 5201 ----CCTGAGAGAGCCATGCTCCATAATCCAAACCAAGCTCAAGGT----- 5242
QY 1334 rLeuAspMetProserSerGluuArgLysLysAlaLeuGlySerLeuGlyGlySerProse 1354
Db 5243 ----GAGAGTCTCTGTCGCAAAATGAGGGGACACACAGTA-----CAGCACCCCGA 5286
QY 1354 rLeuProaAspGlnAspLysAlaGlyCysIleGluValGlnValGlnSerLeuAspThrTh 1374
Db 5287 AGCCCAACAGAAAGAAAG-----CAGAGTGAAGAAACCCCATTCAC 5328
QY 1374 rValThrgInThraAla-----GluAlaValGluLysValIleGly 1387
Db 5329 TCCTCTCTGATGATGACTTCTGACCTTAAGCAAGATTCCCTCCAGAGAGATTGCTCCA 5388
QY 1387 uThrValAlaIleSerGlyThrgLysLysSerProgluGlyValGlyAlaHIsLeu----- 1405
Db 5389 AGAAATCACTGTTGAGAAAGAGATCCCAACCAAGAGATGTCGCCCAAGACCTTCCGCC 5448
QY 1406 ----LeuProAlaGluLysSerSerAlaThrgLysIleStrpThrLeuGlnHIsAl 1423
Db 5449 ACCTCCCAAGCAGACCCGATGATGAGAGCTCAAGCAGAGTTGAGGTTGATTCAT 5508
QY 1423 aGluAspThrValProLeuGlyProGluSerGlnAlaGluSerIleProIleIleValTh 1443
Db 5509 CATTGAAAGTGAAGCTGATGACCCACCCAGCGAT---CCAGATCCCAATCCCAACT 5565
QY 1443 rProAla-----ProGluSe 1448
Db 5566 GCCTTCTGTAAGTGAAGCAAGCTTCACCTCTGCTGCGCTCGGGGGATCCACACCA 5625
QY 1448 rThrLeuHIsProaAspLeuGlnGlyLysSerAlaSerGlnArgLysArgserGluGly 1468
Db 5626 GAGCCCCCTACTAAGTGAACAGATGATCAACAGGACAGAGGAGGAGGCTGAGTC 5685
QY 1468 uGluAspLysProaAspAlaGlyProaAspAlaAspGlyLysGlu-----SerTh 1484
Db 5686 TACTCCATCTCCAGCTCTTCCGCCAGACAAAGGCTGTGATGTGACACCAAGCTTCCAG 5745
QY 1484 rAlaIleGluLysValLeuLysAlaGluProGluIleLeuGluLeuGluSerLysSerAla 1504
Db 5746 CACCTGAGAGAGATTCCT---ATGAGACCCCAAGTATGTCTGCCACCAAGTGTCACTTC 5802
QY 1504 nLysIleValLeuAsnValIleGlnThraAlaValAspGlnPheAlaArgThrgLuthraI 1524
Db 5803 CACAAGTGCACACAGCCATTCGACAGGCTGTC-----AGTCTGTG 5844
QY 1524 aProGluThrHIsAlaTyraAspserGlnThrgLInValProAlaCysArgLeuAspserAl 1544
Db 5845 CCCTTCCCTCAAT-----GAGGCCCCCCCCCCCCCAATTAACCTTA 5886
QY 1544 gGluProaAspArgCysTrpThrLysMetLysAspAlaLysMetLysHIsProValProgl 1564
Db 5887 AAAGCCT-----TTAGAAAGAAACACGACCTCACTGACAAA 5925
QY 1564 nProArgLysAspLeuGlnValLeuThraValLeuGluAlaTrpAlaGlnProArgLysCy 1584
Db 5926 C---AACTCGAATATACAAAGCCTCGAGGTCGTGTAGCT---GCTGACAAAGAAAGGT 5979
QY 1584 sLeuProArgLeuGlnLeuLys-----AlaProValSer 1595

```

```

Db 5980 GCCTCAGTCAATGCTCCCAAAATTAACCTCTGTATTAGCCGAGATGCTGTACGC 6034
RESULT 9
US-10-152-319A-1780
; Sequence 1780, Application US/10152319A
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Higgs, Brandon
; APPLICANT: Castle, Arthur
; APPLICANT: Elashoff, Michael
; TITLE OR INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5089-US
; CURRENT APPLICATION NUMBER: US/10/152,319A
; PRIOR FILING DATE: 2002-05-22
; PRIOR APPLICATION NUMBER: US 60/292,335
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/297,523
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,925
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,810
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/303,807
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/303,808
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/315,047
; PRIOR FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: US 60/324,928
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/330,867
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: US 60/330,462
; PRIOR FILING DATE: 2001-10-22
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2221
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1780
; LENGTH: 15231
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. NM_022401
US-10-152-319A-1780
Alignment Scores:
Pred. No.: 1,06e-09 Length: 15231
Score: 427.00 Matches: 366
Percent Similarity: 33.95% Conservative: 258
Best Local Similarity: 19.91% Mismatches: 745
Query Match: 5.29% Indels: 471
DB: 6 Gaps: 70
US-09-902-432-4 (1-1596) x US-10-152-319A-1780 (1-15231)
QY 3 AlAGlySerSerThrgLInLysArgserProgluGlnProAlaGlySerAspThrProser 22
Db 4467 GCTGGCTGAGCAACAGCGGCGCA-----GAGCCGAGAGCGCTGGC 4508
QY 23 GluLeuValLeuSerGlyHIsGlyProAlaAlaGluAlaSerGlyAlaAlaGlyAspPro 42
Db 4509 CGA-----GTCGAGAGCGCGCTGGAAGACAGCGGAGCTGTGAGAGGCCA 4556
QY 43 -AlaAspAlaAspProAlaThrLysLeuProGluLysAsnGlyGlnLeuSerSerValAs 62
Db 4557 TCCCAAGCGCAAGGACAGAGCCGAGCTGGAGGACAGAACTGCAAGCGCGCATGCAGGA 4616
QY 62 nGlyValAla-----GluGlnGlyAspValHIsValaGlnGluGluAsnGlnGluGly 80
Db 4617 GGAGGTGAGCGCGCGGAGAGAGCGCGGTGAGCCACACAGCAACAGAAAGCGACATCCA 4676

```

QY 80 nGluGlu-----GluValValAs 86
| | | | |
Db 4677 AGAGGAGCTGCAGCATCTGCGCAAGCTCAGAGGCAGAGATCCAGGCCAAGGCCAGCA 4736

QY 86 pGluAspValGlyGluArgGluSerGluAspValArgGluAspArgValGluGluMe 106
| | | | |
Db 4737 GGTGGAGGCTGCAGAGCGCAGCGCATTTGAGGAAGAGATCCGCGTAGTCCGTCT 4796

QY 106 tAlaAlaAsnSerThrAlaValGluAspIleThrLysAspGlyGlnGluThrSerG1 126
| | | | |
Db 4797 GCAGCTAGACAACT-----GAGCGTCAGCGTGGAGGGCGGAGGATGAGTGCA 4847

QY 126 uIleIle-----GluGlnIleProAlaSerGluAsnValGluGluMetVa 142
| | | | |
Db 4848 GGCTCTCGTGCCAGCGGCTCAGGAGGCGAGGAGCAGACAGAGCGCGAGCTCAGGAGGAAGC 4907

QY 142 lGlnProAlaGluSerGlnAlaAsnAspValGlyPheLysLysValPheLysPheValG1 162
| | | | |
Db 4908 CGAGCGCTTCGGAGGCGAGGTGCAGGATGAGACCAACGCAACGGCAGCGCGGAGCGCA 4967

QY 162 yPheLysPheThrValLysLysAsp-----LysAsnGluLysSerAspThrValG1 179
| | | | |
Db 4968 GCTGGCCCTCGGTGAAGCCAGAGCGGAGCGGCGGAGAGAGCAGCGCGGCCCTGCA 5027

QY 179 nLeuLeuThrValLysLysAspGluGlyAlaGlu-----AlaLe 194
| | | | |
Db 5028 GGCTCTCGATGAACCTGAACTGAGGCGGAGGAGCGGCGGAGCGGCTGTGCCAAGCCGA 5087

QY 194 rValGlyAlaGlyAspHisGlnGluProSerValGluThrAlaValGly----- 210
| | | | |
Db 5088 GGCAGAGAGGCTCGCAAGTGCAGGTAGCCCTGGAGACAGCGCGTAGTGCAGAAGT 5147

QY 211 -GluSerAlaSerLysGluSerGluLeuLysGlnSerThrGluLysGlnGluGlyThrLe 230
| | | | |
Db 5148 GGAGCTGCAGAGAGCGTCCGTCCTTTGCAGAGAGACCGCACAGTTGAGGCGCAGCT 5207

QY 230 uLysGlnGluGln---SerSerThrGluIleProLeuGlnAlaGlu-----SerAs 246
| | | | |
Db 5208 GCAGGAAGAGCACGTGACAGTGACACAGCTGCGGAGGAGGCGGAACGGCGGCACAGCA 5267

QY 246 pGlnAlaAlaGluGluGluAlaLysAspGluGlyGluGluGlnGluLysGluProTh 266
| | | | |
Db 5268 GCAGGCTGAAGCCGAGCGAGCCCGTGAAGAACCGGAGCGGAGCTGGAGCGCTGGCAGCT 5327

QY 266 rLysSerProGluSerProSerProValAsnSerGluThrThrSerSerPheLysLy 286
| | | | |
Db 5328 GAAGGCCAATGAGCGCTGCGGCTGCGGCTGCAGGCAGAGAGGAGTGCCACAGCAGAGAG 5387

QY 286 sPhePheThrHisGlyTrpAlaGlyTrpArgLysLysThrSerPheLysLysSerLysG1 306
| | | | |
Db 5388 CCTG-----GCCAGCGCGATGCGGAGAAGCAGGAAGGA 5420

QY 306 uAspAspLeuGluThrAlaGluLysArgLysGluGlnGluAlaGluLysValAspGluG1 326
| | | | |
Db 5421 AGAGCGCAACCGGAAGCCCGCGGCGGCAAGGCAGAGAGCAGCGCCGCTGGCGGAGCG 5480

QY 326 uGluLysGluLysThrGluProAlaSerGluGluGlnGluProAlaGluAspThrAspG1 346
| | | | |
Db 5481 AGAGCTGGCTGAGCAGAG---CTGGAGAGAGCAGCGCGCATCAGAGGGCCCGCCCA 5537

QY 346 nAlaArgLeuSerAlaAspTyrrGluLysValGluLeuProLeuGluAspGlnValGlyAs 366
| | | | |
Db 5538 GCAGCGCTGGCTGCCAGCAGAGGCTGATTGCGCTCGCGGCGAGAGCAGCGAGGAGGTGA 5597

QY 366 pLeuGluAlaSer---SerGluGluLysCysAlaProLeu-----AlaThrGluVa 382
| | | | |
Db 5598 GCATACGCGGAGCTGTGGAGGAAGAGCTGGCCGCGCTACAGCACCAAGCGCAGCAGC 5657

QY 382 lPheAspGluLysMetGluAlaHisGlnGlu-----ValValAlaGluValHisVa 399
| | | | |
Db 5658 CACACAGAGCGCGAGGAGCTGGAGGCTGAGCTGGCGAAGGTTTCGGGCAGAGATGGAGGT 5717

QY 399 lSerThrValGluLysThrGluGluGlnGlnGlyGlyGlyGlyGluAlaGluGlyGlyVa 419
| | | | |
Db 5718 ACTGTGGCCAGCAAGGCAGAGCCGGAAGGAGGTCTCGCTCCACAGTGAA----- 5769

QY 419 lValValGluGlyThrGlyGluSerLeuProGluLysLeuAlaGluProGlnGluVa 439
| | | | |
Db 5770 -----AAGTCCAAGCAGAGGCT 5786

QY 439 lProGlnGluAla-----GluProAlaGluGluLeuMetLysSerArgGluMe 455
| | | | |
Db 5787 GGAAGCTGAGGCGAGCGCGGTTCGAGAGCTGCTGAGGAGGCTGCCCGCTCGCTCTCT 5846

QY 455 tCysValSerGlyGlyAspHisThrGlnLeuThrAspLeuSerProGluGluLysThrLe 475
| | | | |
Db 5847 GGCAGGAGGAGCAAGCGGCACCGGAGTTGCC-----GAGGAGACCGCGC 5894

QY 475 uProLysHisProGlu-----GlyIleValSerGlu-----ValGluMetLeuSerSerG1 492
| | | | |
Db 5895 ACGCCAGCGCGCGAGCGGAGCGAGGTGCTTACGGAGAGCTGGCTGCCATCAGTGAGGC 5954

QY 492 nGluArgIleLys-----ValGlnGlySerPr 501
| | | | |
Db 5955 CACAAGGCTCAAGACGAGCGAGAGATTGCACCTCAAAGAGAGAGGCGCGAGAACGAGCG 6014

QY 501 oLeuLysLysLeuPheSerSerSerGlyLeuLysLysLysLeuSerGlyLysLysGlnLysG1 521
| | | | |
Db 6015 CTTGAGCGCCTGGCTGAAGATCAGGCTTCAGCGCGCGCGCTGGAGGAGCAGGAGCAGC 6074

QY 521 yLysArgGlyGlyGlyAspGluGluProGlyGluTyrrGlnHisIleHisThrGluSe 541
| | | | |
Db 6075 ACAGCAAGCGAGACATAGAGGCGCTGGCCAG-----CTGCCAAGGCATC 6125

QY 541 rProGluSerAlaAspGluGlnLysGly-----GluSerSerAlaSerSerProGluG1 559
| | | | |
Db 6126 CGAGACGAGCTGGAGCGACAGAGGGGTGGTGGAGGATACCTCCGCGAGCGCGCGCA 6185

QY 559 uProGluGluThrThrCysLeuGluLysGlyProLeuGluAlaProGlnAspGlyGluAl 579
| | | | |
Db 6186 GGTGGAGGAGGAGATCATGGCTCTGAAGCGAGCTTCGAGAGAGCGCGCGCTGCCAAGGC 6245

QY 579 aGlu-----GluGlyThrThrSerAspGlyGluLysLysAr 591
| | | | |
Db 6246 AGAACTGGAGCTGGAGCTTGGCGGCATCCGAGCAATGCCAGAGACCATCGCGCAGCAA 6305

QY 591 gGluGlyIleThrProTrpAlaSerPheLysLysMetValThrProLysLysArgValAr 611
| | | | |
Db 6306 GGAGCTGCGCGAGCAGGAGCGCGCGCGCGGCGGAGTTGGCAGCTGAGGAGGAGCAGAG 6365

QY 611 gArgProSerGluSerAspLys-----GluGluGluLeuGluLy 624
| | | | |
Db 6366 GCGCCGGAAGCGGAGGCGGCTGCAGAGAGCGCTTGGCAGCGGAGGAGGAGGAGCGCAGC 6425

QY 624 sValLysSerAlaThrLeuSerSerThrAspSerThrValSerGluMetGlnAspGluVa 644
| | | | |
Db 6426 GCAGCCCAAGTGCACACTGGAGAACTCAGCGGCTCAAGGGCCCAAGGTT---GAGGAAGC 6482

QY 644 lLysThrValGlyGluGluGlnLysProGluGluProLys----- 657
| | | | |
Db 6483 CGCGCGCTTCGAGAGCAGGAGCTGAGCAGAGGTCTGGCAGCGAGCTGCAGCTGCCCGCAGGA 6542

QY 658 -----ArgArgValAspThrSerValSerTrpGluAlaLeuIleCys----- 671
| | | | |
Db 6543 GGCTGCCCAAGAAACGCTGCAGCGGAGGAGGAAGGGCGCAGCGCTTTGTTGTCGAGCAGCG 6602

QY 672 -----ValGlySe 674
| | | | |
Db 6603 AGAAGAGGAGCTGCAGCAGACTCTTCAGCAAGAGCAGCAACATGCTGGAGCGGCTGGCGAG 6662

QY 674 rSerLysLysArgAlaArgLysAlaSerSerSerAspAspGluGly-----GlyPr 691
| | | | |
Db 6663 CGAGGCGAGGAGCGCGCGGAGCTCTCAGGAGCGCGGAGGAGCGCGGAGCAGCAGCAGGA 6722

QY 691 oArgThrLeuGlyGlyAspSerHisArgAlaGluGluAlaSerLysAspLysGluAlaG1 711


```

Db 6723 ACCTGAGCGACGGCCTTACGAGAGCAAGTGGAGAGCGCCGAGCGCTGAGAGCTGCG-- 6780
Qy 711 yThrAspAlaValProAlaSerThrGlnGlnAspGlnAlaGlnGlySerSerSerPr 731
Db 6781 -----GCGAGAGGACGAGGCTCAGCGCCAGCGCCAGCGCAGCGCGCTCG 6824
Qy 731 oGluProAlaGlySerProSerGlnGlnGlnValSerThrTTPGluSerPheLysAr 751
Db 6825 AGAG----- 6828
Qy 751 gLeuValThrProArgLysLysSerLysSerLysLeuGlnGlnLysAlaGluAspSerSe 771
Db 6829 -----AAACTGCGCAAGAGAGCGGAGCGGAGCGCGCGCTGCGGCCCA 6872
Qy 771 rValGlnLeuSerThrGlnLleGluProSerArGlnGlnGlnSerTTPValSerLleL 791
Db 6873 GCGCGAGCGCGCTGTAACAGAGAGCGGAGCGCGGAGAGTGAAGAGACACA 6932
Qy 791 sLysPheLleProGlyArGArgLysLysArgAlaAspGlyLysGlnGlnAlaThrVa 811
Db 6933 GAAGTTTGACAGACAGACCTACGCGCAAGAGCTCAGGTAGAGACAGAGCTGACACGCT 6992
Qy 811 lGluAspSerGlyProValGlnLleAsnGluAspAspProAsnValProAlaValAlPr 831
Db 6993 AGAG-----CTGACGCTCGAGAGAGACCGAC--CACACAGAGACATCTCGGA 7037
Qy 831 oLeuSerGluTyArAsnAlaValGluArgLysLysMetGlu--AlaGlnGlyAsnThrGl 850
Db 7038 T---GAGGAGCTGCACGCGCTAAAGCTGAGTAAACAGAGCGCGCGGAGCGTGA 7094
Qy 850 uLeuProGlnLeuGlnValAlaValTyValSer---GluGlnLeuSerLysThrLeuVa 869
Db 7095 GGTAGAGGAGAGCTCTCTCTCTGCGCGGAGATGAGAGAGAGCTGCGCAAA----- 7146
Qy 869 lHisThrValSerValAlaValLleAspGlyThrArgAlaValThrSerValGlnGlu 889
Db 7147 -----CTCAAGGCTCGCATTAAGCTGAAACCG 7175
Qy 889 gSerProSerTTPLeSerAlaSerValThrGluPro---LeuGlnHisThrLleGlnGl 908
Db 7176 GCGACATCTCTCGTGCAGACAGAACACACAGCGCTCTGAGAGAGAGCGCAGAA 7235
Qy 908 uAlaMetProProValGlnGlnValThrGlnLysAspLleLleAlGlnGlnThrProVa 928
Db 7236 GATGAACAAGGTGGCAGAGAGCTGACGCGTGAAGCTGAGCCGAGAGAGCAGCAAG 7295
Qy 928 lLeuThrGln-----ThrLeuProGlnGlnLysAspAlaHisAspAspMetVa 944
Db 7296 GCTGCGGCGAGTACCGCAGAGAGAGCTGCGCCGAGCGCGGCGCTGCGGAGAGATGCT 7355
Qy 944 lThrSerGluValAspPheThrSerGlnAlaValThrAlaThrGlnThrSerGlnAlaLe 964
Db 7356 GAAGGAGAAAGATCCAGCGGTGACAGAGACCAAGGCTCAAGCGTGAAGCTGAGCTGCT 7415
Qy 964 uArgThrGlnGlnValThrGlnLleAspGlyAlaGlnGlnThrThrAspMetValSerAl 984
Db 7416 GCGAGCAGAGAGAGAGCTG-----GCACAGAGAGAGAGCGCGCGGCTGACAGCG 7463
Qy 984 aValSerGlnLeuThrAspSerProAspThrThrGlnGlnAlaThrProValGlnGlnVa 1004
Db 7464 G-----GACAAAGAGCAAAATGGCTCAGACAGTGGTGAAGGA 7499
Qy 1004 lGluSerGlyValLleuAspThrGlnGlnGlnGlnAlaArgGlnThrGlnAlaLleLeuGlnAl 1024
Db 7500 GACACAGGCTTCCAGCGAGACCTGAGAGGTGAGCGGACGCGCAGCTAGAAATAGAGCG 7559
Qy 1024 aValAlaAspLysValLys----- 1030
Db 7560 AGAGGCTGAAGCGCTCAAGTTGGCATGGCTAGATGATGACCGGCGTCAAGCGCGTGCAGA 7619
Qy 1031 -GluGluSerGln----- 1034

```

```

Db 7620 GGAGGATGGCCAGCGCTTCGGAACAGAGCTGAAGAGATTCGCGCAAAAGCTCACCGCAC 7679
Qy 1035 -----ValProAlaThrGlnThr-----ValGlnArgThrGl 1045
Db 7680 TGAATCGCTACACAGAGAAAGTGATGATGGAGACTCTCGAGATTCAGCGACAGACA 7739
Qy 1045 ySer-----LysAlaLeuGlnLysValGlnGlnValGlnGlnLysAspSerGl 1060
Db 7740 GAGTACCAAGATGCCAGCGCTGTAGGAGCGCCATTCTGAGCTGGAGCGTGAAGAGA 7799
Qy 1060 u-----ValLeuAlaSerGlnLysGlnLysAspValMe 1071
Db 7800 GAAGCTCAAGCAGAGAGGGAATTACTGACGTCAAGTCTGAGAGAGATGACAGACTGTGCA 7859
Qy 1071 rProLysGlyProValGlnGlnAlaGlyAlaGlnHisLeuAlaGlnGlnLysSerGlnThrGl 1091
Db 7860 GCAGAGCAGATACGTGAGAGAGACACAGCGCTCGAGAAAGCTTCTCTGAGAAAGGA 7919
Qy 1091 yGlnAlaThrProGlnLysLeuGlnValProGluValThrAlaAspValAspHisValAl 1111
Db 7920 CAGCTTGCTCAACCGCAACGCTTCATGAGAGAGAGAGAGAGCTGAGACAGCTTTT 7979
Qy 1111 aThrCysGlnValLleLysLeuGlnGlnLeuMetGlu-----GlnAl 1125
Db 7980 CCAGAGACGAGCTGGCAAAAGCAAAACACTGTGACAGAGAGACAGCGCGGACAGACAGA 8039
Qy 1125 aValAlaProGluSerSerGluThrLeuThrAspSerGlnThrAsnGlySerThrProLe 1145
Db 8040 GATGAGACAGAAAGACAGAGAGCTGTGAGCAGCATGAG-----GAGGCCCGAG 8090
Qy 1145 uAlaAspSerAspThrAlaAspGlyThr----- 1154
Db 8091 GCGGACGCTGAGGAGAGAGGAGGTGAGAGCGCAAGAGAGAGAACTGACGCTGTGGA 8150
Qy 1155 -----GlnGlnAspGluThrLleAspSerGlnAspSerLysAlaThrAlaAlaVa 1171
Db 8151 GCAGACGCGCGCAGCAGCAGAGAAACTACTGCGACAGAGAGAACAGAGAGCTCGGAGCG 8210
Qy 1171 lArgGlnSerGlnValThrGlnGlnGlnAlaAlaThrAlaGlnLysGlnGlnProSerTh 1191
Db 8211 GCTGAGCGCTGAGAGAAAGCACCGAGCTGCTGTCGCGCACTGTGAGAGATGCGCAC 8270
Qy 1191 r-----LeuProAsnAsnValProAlaGlnGlnGlnHisGl 1203
Db 8271 CTCCAGGCTGCTGCCACAAAAGCACTGCCAAT----- 8304
Qy 1203 yGlnGluProGlyArGAspValLleuGluProThrGlnGlnGlnLysLeuThrAlaAlaAlaVa 1223
Db 8305 -----GCGCGCAGCGCACTTGATGCCCTCCATGAG----- 8337
Qy 1223 lProValLeuAlaLysThrGlnValGlyGlnGlnGlnValAspTTPLeuAspGlyGl 1243
Db 8338 -----GCCAGCCGAGTACACTTTGAGGAGATTACGTCAAGAGGTCCAGCTCA 8387
Qy 1243 uLysValLysGlu-----GluGlnGlnValPheValHisSerGlyProAsnSe 1259
Db 8388 GCAGCTACAGAGAGAGAGCATTTGAGCATGAGAGAACTGACGCTTGACACAGGGTCA 8447
Qy 1259 rGlnLysAlaAlaAspValThrTyAspSerGlnVal-----Me 1272
Db 8448 CACCAAGGTGCTGAGCTCAGCCAGCGAGAGATGAGGCCACTGACAGGCGGAG 8507
Qy 1272 rGlnValAlaGlyCysGlnGlnLysGlnSerThrGlnValGlnSer----- 1287
Db 8508 CAGATCGCAGAGATTGCTCTGAAGCCACCAATAGAAACTGAGTGTCTACACAGCCCT 8567
Qy 1288 -----LeuSerLeuGlnGlnGlnGlnMetGlnThr 1297
Db 8568 ACAGCGGAGCTGCTCAGCCCTGGAACAGCCCTTATCTTACTTGAAGCCGACAGCCCTC 8627
Qy 1297 r-----AspValGlnLysGlnLysArgGlnThrLysProGlnGlnValSerGl 1313
Db 8628 GGGCTTCTGCTGAGACCTGTCTCGAACCAGCGGCTGACGCTGATATGAGCTGTGAAGGA 8687

```

```

QY 1313 u----- 1313
DB 8688 GGGTGTGGTCCGAGCTGCACCACAAAGCTGCTCAGCTGAGCGTGCCGCTCACTGG 8747
QY 1314 -----GluGlyGluGlnGluThrAlaAlaProGluHisGlu----- 1325
DB 8748 CTACAGGACCCTTACACAGGAGACAGAT-CTCTCTCTTCCAGGCGCATGAAGAAGACC 8806
QY 1326 -----GlyThrTyrglyLysProValLeuThrLeuAsp----- 1336
DB 8807 TCATTGTCAGGACCATGGCTCGCTGCTGGAGCCAGATGCCACAGGTGGCATCA 8866
QY 1337 -----MetProSerSerGluArgGlyLysAlaLeuGlySerLeu----- 1349
DB 8867 TTGACCCCTGTACACAGCACCGGTGTTCCGCTGGACGGCTTACACAGCGTGCTTCTCG 8926
QY 1350 -----GlyGlySerProSer-Leu-----ProAspGlnAspLysAlaGlyCysIleGluV 1366
DB 8927 ATGAGGAGATGAACCGGTGCTGGCTGACCCAGCGATGACACCAAGGGCTCTTTGACC 8986
QY 1366 alGlnValGlnSerLeuAspThrThrValThrGlnThrAlaGluAlaValGluValI 1386
DB 8987 CCAACTCTAC-----GAGAACTCA 9007
QY 1386 leGluThrValIleSerGluThrGlyGluSerProGluCysValGlyAlaHisLeuL 1406
DB 9008 CGTACCTGTGACGCTGCTGGAGCGCTGTGTGGAGGACCCCGAG---ACAGGCGTGGCTCC 9064
QY 1406 euProAlaGluLysSerSerAlaThrGlyHis-Trip---ThrLeuGlnHisAlaGlu 1424
DB 9065 TGGCACTCTACAGACAGAGTGGCCAAAGGTGGTGTGAGTGTGTACATGACACCGAGGCC 9124
QY 1425 AspThrValProLeuGlyProGluSerGlnAlaGluSerIleProIleValThrPro 1444
DB 9125 GTGACGCTCTCGAAAGGCCACAGTGTCTGCACCATTCGCGCAGTTCACAGGCGAAGCCG 9184
QY 1445 AlaProGluSerThrLeuHisProAspLeuGlnGlyLysSerAlaSerGlnArgGlu 1464
DB 9185 TGACCATCTGGGAGA-----TCATCACTCAGAGTACT 9217
QY 1465 ArgSerGluGluGluAspLysProAspAlaGlyProAspAlaAspGlyLysGluSerThr 1484
DB 9218 TCACAGCGGAGCAGCAGCGGACCTGCTCCGCGCAGTTCGCCACG-GGCCGC-----ATC 9270
QY 1485 AlaIleGluLysValLeuLysAlaGluProGluIleLeuGluLeuGluSerLysSerAsn 1504
DB 9271 ACGGTGGAGAGATCATCAGATTGTCATCAGGTGTGTAGAGAAACACAGCGCGAAGGGC 9330
QY 1505 LysIle-ValLeuAsnValIleGlnThrAlaValAspGlnPheAlaArgThrGluThrAl 1524
DB 9331 CAGCTCTGCTTTG----- 9343
QY 1524 aProGluThrHisAlaTyraSerGlnThrGlnValProAlaCysArgLeuAspSerAr 1544
DB 9344 -----AGGCGCTCCGTCCTTGTGCTGTGCAGAGC-- 9376
QY 1544 gGluProAsnArgCysTrpThrLysMetLysAspAlaLysMetLys 1559
DB 9377 -----TGCTGGACA---GTGGAGTCATCATGTATGAAG 9406

```

RESULT 10

```

US-09-724-676-25032
; Sequence 25032, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 25032

```

```

; LENGTH: 14335
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676-25032

Alignment Scores:
Pred. No.: 2,38e-09 Length: 14335
Score: 418.00 Matches: 334
Percent Similarity: 35.65% Conservative: 255
Best Local Similarity: 20.22% Mismatches: 701
Query Match: 5.18% Indels: 364
DB: Gaps: 61

```

```

US-09-902-432-4 (1-1596) x US-09-724-676-25032 (1-14335)
QY 5 SerSerThrGluGlnArgSerProGlu-GlnProAlaGlySerAspThrProSerGluLe 24
DB 4187 AGCAGACGGCGGAGAGAGAGCGGCGGCTGGCCAGGTGGAGGCGCGCTGGAGAAGC 4246
QY 24 uValLeuSerGlyHisGly-----ProAlaAlaGluAlaSerGlyAlaAlaGlyAspPr 42
DB 4247 AGCGGCAGCTGGCCGAGGCGCACGCCCAAGGCACAGGCGGAGCGGAGCGAAGG 4306
QY 42 oAlaAspAlaAspProAlaThrLysLeuProGlnLysAsnGlyGlnLeuSerSerValAs 62
DB 4307 AGCTGCAGCAGCGCATGCA-----CGAGGAGTGTGTGCGCGGAGGAGCGCG 4354
QY 62 nGly-ValAlaGluGlnGlyAspValHisValGlnGluGlnGluGlnGluGlnGluG 82
DB 4355 CGGTGACCGCGCAGCAGCAGAGCGGCGCATTCAGAGGAGGTGCAGCAGCTCGCGCAGA 4414
QY 82 luGluValValAspGluAspValGlyGlnArgLysGlu----- 95
DB 4415 GCTCGGAGCGGAGATCCAGGCCCAAGGCCGAGGCGGAGGCGGCTGAGCGAGCGCGC 4474
QY 96 --AspValArgGluLysAspArgValGluGluMetAlaAlaAsnSerThrAlaValGluA 115
DB 4475 TGGCATCATGAGGAGGAGATCCGCGTGTGCGCTGCGATTTGGAGGCGCACCGAGCGCAGC 4534
QY 115 spileThrLysAspGlyGlnGluGluThrSerGluLeu---lleGluGlnIleProAlas 134
DB 4535 GTGGCGGGCTGAGGGGAGCTGCGGCGCTGCGTGCACGGCGGAGGAGGCTGAGGCAC 4594
QY 134 erGluAsnAsnValGluGluMetValGlnProAlaGluSerGlnAlaAsnAspValGlyP 154
DB 4595 AAAAGCGACAGCGCAGGAGGAGCGCGCTGCGGAGGCGAGGTGCAGCAGCAGAGCC 4654
QY 154 heLysLysValPheLysPheValGlyPheLysPheThrValLysLysAsp-----L 171
DB 4655 AGCGTAAGCGGAGCGGCGGAGGTGGAGTGGCTGCGCGCTGAAGGCGGAGCGAGCGCGG 4714
QY 171 ysAsnGluLysSerAspThrValGlnLeuThrValLysLysAspGluGlyGluGlyA 191
DB 4715 CGCGCGAGAGCAGCGCGCGCTGCGAGGCGCTGAGAGGCTGCGGCTGCGAGGCGGAGG 4774
QY 191 laGlu-----AlaSerValGlyAlaGlyAspHisGlnGluProSerValG 206
DB 4775 CGGAGCGCGCTGCGGCGGCGGAGGTGCGGCGGCGGAGGTACAGGTGCGCTGG 4834
QY 206 luThr-----AlaValGlyGluSerAlaSerLysGluSerGluLeuLysGlnS 222
DB 4835 AGACGCGCAGCGCGCTGCGAGGCGGAGTGCAGAGCAAAAGCGCGCTCTTTCGCGCAGA 4894
QY 222 erThrLysLysGlnGluGlyThrLeuLysGlnGln---SerSerThrGluIleProL 241
DB 4895 AGACGCGCAGCTGGAGCGCTCCCTGAGGAGGAACCTGCTGTGGTGTGGCACAGCTCGGG 4954
QY 241 euGlnAlaGlu-----SerAspGlnAlaAlaGluGluAlaLysAspGluGlyG 258
DB 4955 AGGAGGCTGAGCGGCGGCGGCGCAGCAGCAGCGCGCGGCGGCGGCGGCGGAGGCGCAG 5014
QY 258 luGluLysGlnGluLysGluProThrLysSerProGluSerProSerSerProValAsnS 278

```

Db 5015 AGCGGAGCTGAGAGCGCTGCGACGCTCAAGGCCAAGAGGCGCTACGCGCTGCGACG 5074
 Qy 278 eRgluThrThrSerSerPheLysPhePheThrHisgluTyrAlaGluTyrArgLysP 298
 Db 5075 CGAGGAGAGTGGCGACGAGAGAGCGCTG-----GCCG 5107
 Qy 298 yRThrSerTherLysLysSerLysGluAspAspLeuGluThrAlaGluLysArgLysGlu 318
 Db 5108 AGCGCGAGGCTGAGAGCGAGAGAGAGAGCGGAGCGGACGCGCGCGCGCGCGGCGCAAG 5167
 Qy 318 lngluAlaGluLysValAspGluGluGluLysGluLysThrGluProAlaSerGluGlu 338
 Db 5168 CGGAGAGAGAGCGCTCGCGCGAGCGGAGCTGCTGACAAAGG---CTGGAAGACAGC 5224
 Qy 338 lngluProAlaGluLysPheThrAspGluAlaArgLysSerAlaAspTyrGluLysValGlu 358
 Db 5225 GCGAGCTGCGGAGAGCGACCGCGACAGCGCTGCGCGCGCGGAGCGAGTTCGCGC 5284
 Qy 358 euProLeuGluLysPheThrAlaGluLysPheLysGluLysSer---SerGluLysLysCysAlaP 377
 Db 5285 TCAGGCGCGCGAGAGCGAGCGGAGCGAGCGGAGCGAGCTGCTGCGGAGAGAGAGCTGGCCC 5344
 Qy 377 roLeu-----AlaThrGluValPheAspGlu----- 385
 Db 5345 GCGCTGAGCTGAGAGCGCGCTGCGACGCCAAGAAAGCGAGAGCTGGAAGCGGAGCTGG 5404
 Qy 386 --LysMetGluAlaHisGluGluVal-----AlaGluValH 398
 Db 5405 CCAAGGTCGGGCGCGAGATGAGAGTGTCTGCTGCGCAGCAAGCGGAGCGGCTGAGAGAGAT 5464
 Qy 398 lValSerThrValGluLysThrGluGluGluGluGluGluGluGluGluGluGluGluGlu 418
 Db 5465 CCGCGCTCCACCGAGCGAGAGTCCAGACGAGCGCTG-----GAGCGGAGGCGC 5512
 Qy 418 lValValValGluGluGluGluGlu-----SerLeuProProGlu----- 431
 Db 5513 GCGGTCGCGCGAGCTGCGCGAGAGCGCGCGCGCTGCGCTGCGCGCGGAGAGAGCGCA 5572
 Qy 432 -----LysLeuAlaGluProGluGluValProGluGluAlaGluProAlaGlu 448
 Db 5573 AGCGGAGCGGAGCTGCGCGAGAGAGAGCGGCGCGGAGCGGCGCGAG---GCGGAGC 5629
 Qy 448 lLeuMetLysSerArgLysMetCysValSerGluLysAspHisThrGluLeuThrAspL 468
 Db 5630 GGGTCCTTCGAGAGAGCTGCGCGCATCGC-----GAGGCCACCGCGC 5674
 Qy 468 euSerProGluGluLysThrLeuProLysHisProGluGluLysValSerGluValGlu 488
 Db 5675 TCAAGACGAGAGCGGAGATCCGCTCAAGAGAGAGAGCGGAGAGCGCGCTGCGC 5734
 Qy 488 euLeuSerSerGluGlu-----ArgLysValGluGluGluSerProLeu 503
 Db 5735 GCGTCGCGAGAGCGAGCGCTTCAGCGCGCGCGCTGCGAGGAGCGGCGCGCAACACA 5794
 Qy 503 yAluLysLeuHeSerSer-----SerGluLysLysLeuSerGlu-----LysL 518
 Db 5795 AGCGTCGACATCGAGAGCGCGCTGCGCGAGCTGCGCAAGGCGCGCTGCGCAAGCGGAGC 5854
 Qy 518 yGluLysGluLysArgGluGlu-----GlyGluAspGluGlu 531
 Db 5855 GCGAGAGAGGCGCT-GGTGAGAGAGCGCTGAGAGAGCGGCGCGGAGTGGAGAGAGATC 5913
 Qy 531 roGluLysGluLysHisLeuHisThrGluSerProGluSerAlaAspGluGlu-----L 549
 Db 5914 CTGCGC-----GCTGAAGCGGAGCTTCGAGAGAGCGCGCGCTGCGCAAGCGGAGC 5961
 Qy 549 yGluLysSerSerAlaSerSerProGluGluProGluGluThr---ThrCysLeuGlu 567
 Db 5962 CTGAGACTGAGAGCTGGAGCATCCGACAGCAAGCGAGAGACACGCTGCGCAAGAGAG 6021
 Qy 568 LysGluProLeuGluAlaProGluAsp----- 576
 Db 6022 CAGGCGGAGCTGAGGCGCGGAGCGAGCGGAGCTGCGCGGAGAGAGAGCGGCGCGC 6081

Qy 577 GlyGluAlaGluGluGluGluThrThrSerAsp-----GlyGluLys 589
 Db 6082 CGTGAGCTGAGAGCGCGCTCGACAGAGCGCTGCGCGCGCGAGAGAGCGCGACGCGAG 6141
 Qy 590 LysArgGluGluLysLeuThrProThrAlaSerPheLysLysMetValThrProLysArg 609
 Db 6142 CGGAGCGCGCGCTGAGAGAGCTGAGCGCGTGAAGCCAACTGAGAGAGCGCGCGC 6201
 Qy 610 ValArgArgProSerGluSerAspLysGluGluGluGluLysValLysSerAlaThr 629
 Db 6202 CTGCGGAGCGAGCGAGACGAGAGCTGCGCGCGAGCTGCGAGCTGCGCGAGAGCGCGC 6261
 Qy 630 LeuSerSerThrAspSerThrValSerGluMetGluAspGluValLysThrValGlu 649
 Db 6262 CAGAAAGCGCTG-----CAGCGGAGAGAGAGCGACAGCGCTTCCGCGTG 6306
 Qy 650 GluGluLysProGluGluProLysArgArgValAspThrSerValSerTrpGluAlaLeu 669
 Db 6307 CAGCAGAGAGAGAGAGAGCTTCAGCAGACGCTGCGAGAGAGAGAGAGC-----CTGCTG 6360
 Qy 670 lLysValGlySerSerLysLysArgAlaArgLysAlaSerSerSerAspAspGlu 689
 Db 6361 GACCACTGCGCGCGAGCGGAGCGCGCGCGCGCGCTGAGAGAGCGGAGAGAGCGC 6420
 Qy 690 -----GlyProArgThrLeuGluLysAspSerHisArgAlaGluGluAlaSerLys 706
 Db 6421 CCGGTGAGCGGAGCGGAGCGGCGCGAGCGCGCGCGAGGTGAAGAGCGCGAGCGG 6480
 Qy 707 AspLysGluAlaGluThrAspAlaValProAlaSerThrGluGluGluAlaGlu 726
 Db 6481 CTGAAGCACTG-----GCAAGAGAGAGAGAGAGAGCGCGCGCTGAG 6522
 Qy 727 GlySerSerSerProGluProAlaGlySerProSerGluGluGluValSerThrTrp 746
 Db 6523 GCACAGCGCGCTGAGAG----- 6540
 Qy 747 GluSerPheLysArgLeuValThrProArgLysLysSerLysSerLysLeuGluGluLys 766
 Db 6541 -----AAGCTGCCAAGAGAGCGGAGCGGAGAGAGCGG 6570
 Qy 767 AlaGluAspSerSerValGluGluLeuSerThrGluLeuProSerArgGluLys 786
 Db 6571 GCGCGCGCGGAGAGCGGAGAGCGCGCGCTGCGCGAGAGAGAGAGAGAGAGAGAG 6630
 Qy 787 TrpValSerLysLysPheLeuProGluArgArgLysLysArgAlaAspLysGlu 806
 Db 6631 ATGAGAGAGCATTAAGAAATTCGCCAGAGAGAGCTGCGGAGAGAGAGAGAGAGAG 6690
 Qy 807 GluGluAlaThrValGluAspSerLysProValGluLeuAspGluAspProAspVal 826
 Db 6691 GAGCTGACAACTGCGC-----CTGACGCTGAGAGAGAGAGAGAGAGAGAGAG 6729
 Qy 827 ProAlaValProLeuSerGluTyrAsnAlaValGluArgGlu-----Lys 842
 Db 6730 -----CACAGAGAGAGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6771
 Qy 843 MetGlu-----AlaGluLysThrGluLeuProGluLeuGluGluValAla 858
 Db 6772 GCGGAGCGGAG 6831
 Qy 859 TyrValSer---GluGluLeuSerLysThrLeuValHisThrValSerValAlaValIle 877
 Db 6832 CCGCTGACAGATGAGAGAGCTGAGAG----- 6858
 Qy 878 AspGluThrArgAlaValThrSerValGluGluArgSerProSerTrpIleSerAlaSer 897
 Db 6855 -----CTCAAG 6912
 Qy 898 ValThrGluProLeuGluHisThrAlaGluGluAlaMetProProVal---GluGluVal 916
 Db 6913 AATACGAGCGCTTCCTGAG 6972

QY 917 ThrGluLysAspIleIleAlaGluThrProValLeuThrGln-----Thr 932
 Db 6973 CGCGCGGTGAGTGGCGGCGCCAGAGCTCGCGAGTGGCGGAGAGGAC 7032
 QY 933 LeuProGluLysAspAlaHisAspMetValThrSerGluValAspPheThrSer 952
 Db 7033 CTGGCAGCAGCGCGGCTTGGCAGAGAAGTGTCAAGAGAAGATGCGCGGCTGCAG 7092
 QY 953 GluAlaValThrAlaThrGluThrSerGluAlaLeuArgThrGluValThrGluAla 972
 Db 7093 GAGGCCACGGGACTCAAGGCTGAGCGGAACTGCTGCAGCAGCAGAGAAGAGCTTGCAG 7152
 QY 973 SerGlyAlaGluThrThrAspMetValSerAlaValSerGlnLeuThrAspSerPro 992
 Db 7153 GAGCAGCGCGGCGGTGCGAGGAG-----7176
 QY 993 AspThrThrGluGluAlaThrProValGlnGluValGluSerGlyValLeuAspThrGlu 1012
 Db 7177 GACAGGAGCAGATGGCGCAGCAGCTGGCGGAGGAGAGCGGCTTCCAGCGGAGCGTG 7236
 QY 1013 GluGluGluArgGlnThrGlnAlaIleLeuGlnAlaValAlaAspLysValLys-----1030
 Db 7237 GAGCGCAGCGGCGCAGCTGAGATGAGCTGAGCGCTGAGCGCTCAAGCTGCGT 7296
 QY 1031 -----GluGluSerGlnValProAlaThrGlnThrValGlnArgThrGlySer 1046
 Db 7297 GTGGCCGAGATGAGCGGAGCGCCAGCGCTGAGGAGCAGCGCTTCCGGAAG 7356
 QY 1047 LysAlaLeuGluLysValGluGluValGluLysAspSerGluValLeuAlaSerGluLys 1066
 Db 7357 CAGCGGAGGAGATCGGTGAGAGCTGCACCGCAGGAG-----CTCGCCACCCAGGAG 7410
 QY 1067 GluLysAspValMetProLysGlyProValGlnGluAlaGlyAlaGluHisLeuAlaGln 1086
 Db 7411 AAGTGACCTGTGTGACAGACTGAGATCCAGCAGCAGCAGAGTGACCATGATGCCGAG 7470
 QY 1087 GlySerGluThrGlyGlnAlaThrProGluSerLeuGluValProGluValThrAlaAsp 1106
 Db 7471 -----CGCCTGCGGAGGCGCATCGCTGAG 7494
 QY 1107 ValAspHis-----ValAlaThrCysGlnValIleLysLeuGlnLeu--MetGlu 1123
 Db 7495 CTGGAGCGTGAGAGGAGAGTCCACAGAGGCGCAACTGTGCGAGCTCAAGTCTGAG 7554
 QY 1124 GlnAlaValAlaProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThr 1143
 Db 7555 GAGATGACAGCGGTGCAGCAGGAGCAGCTGCTCAGCAGCAGCAGCGGCTGCAGCAAGC 7614
 QY 1144 ProLeuAlaAspSerAspThrAlaAspGlyThrGlnGlnAspGluThrIleAspSerGln 1163
 Db 7615 TTCTCTCTGAAAAGGACAGCGCTG-----CTACAGCGGAGCGGCTTCATCGAGCAGGAG 7668
 QY 1164 AspSerLysAlaThrAlaAlaValArgGlnSerGlnValThrGluGluGluAlaAlaThr 1183
 Db 7669 AAGCCCAAGCTG-----GAGCAGCTTTCAGGAGCAGGAGTGCGCAAG 7710
 QY 1184 AlaGlnLysGluGluProSerThrLeuProAsnValProAlaGlnGluGluHisGly 1203
 Db 7711 GCACAGAG-----CTGCGTGAGGAG-----7731
 QY 1204 GluGluProGlyArgAspValLeuGluProThrGlnGlnGluThrAlaAlaVal 1223
 Db 7732 CAGCAGCGGCGCAGCAGCAGATGAGCAGCAAGCGGCGGCTGTTGGTGGCAGCATGGAG 7791
 QY 1224 ProValLeuAlaLysThrGluValGlyGlnGluGly-----GluVal 1237
 Db 7792 GAGCGCGCGCGGCGAGCATGAGCGCGAGGCGGCTGCGCGCAGCAGGAGGAGCTG 7851
 QY 1238 AspTrpLeuAspGlyGluLysValLysGluGluGlnGluValPheValHisSerGlyPro 1257
 Db 7852 CAGCAGCTGAGCAGCAGCGG-----CGCAGCAGCAGGAGCTGTGCTGAGGAGAACCCAG 7908
 QY 1258 AsnSerGlnLysAlaAlaAspValThrTyrAspSerGluValMetGlyValAlaGlyCys 1277

Db 7909 AGCTCGGTGAGCAGGTGCGAGCTCTGGAGGAGCAGCAGCGGCGCGCTGCGCACTCA 7968
 QY 1278 GlnGluLysGluSerThrGluValGlnSerLeuSer-----1289
 Db 7969 GAGGAGGTCACTGCTCGCAGGTGGTGGCCACAAAGACCTGCCAATGGCCGGATGCA 8028
 QY 1290 -----LeuGluGluGlyGluMetGluThrAspValGluLysGluLysArgGlu 1305
 Db 8029 CTTGATGGCCCGCGCAGAGGCGGAGCAGAGCTTCGATGGCTTCCGCGCGGAG 8088
 QY 1306 ThrLysProGluGlnValSerGluGluGlyGluGlnGluThr-----1319
 Db 8089 GTCTCAGCTCAGAGGTGCGAGGAGCGCGCAT--CCTGAGTGGCGGAGAGCTGCGAGCGTT 8147
 QY 1320 ---AlaAlaProGluHisGluGlyThrTyrGlyLysProValLeuThrLeuAspMetPro 1338
 Db 8148 GCGCGAGGCGCACACCGTGA-----8171
 QY 1339 SerSerGluArgGlyLysAlaLeuGlySerLeuGlyGlySerProSerLeuProAspGln 1358
 Db 8172 -----CGAGCTCGCAGCGGCGGAGAGCGTGGCCACTACTCGAGGCGCGCAGCAG 8222
 QY 1359 AspLysAlaGlyCysIleGluValGlnValGlnSerLeuAspThrThrValThrGlnThr 1378
 Db 8223 TATCGCAGGCGTGTCTGTAAGGCCAC-----CAATGAGAA 8258
 QY 1379 AlaGluAlaValGluLysValIleGluThrValValIleSerGluThrGlyGluSerPro 1398
 Db 8259 GCTGAGTGTTTACGCCCGCTCGCAGGAGCAGTGTGAGTCCCGGACCGGCTCATCTCT 8318
 QY 1399 Glu-----1399
 Db 8319 GCTGGAGCGCAGCGCGCTCAGGCTTCTGCTGGAGCCCTGTGCGGAACCGCGGCTGAC 8378
 QY 1400 -----CysValGlyAlaHisLeuLeuProAlaGluLysSer 1411
 Db 8379 CGTCAACGAGGCTGTGAAGAGGCTGTGGTGGCGCCGAGCTGCACCAAGCTGTCTGTC 8438
 QY 1412 SerAlaThrGlyGlyHisThrThrLeuGlnHisAlaGluAspThrValProLeuGlyPro 1431
 Db 8439 GCGCAGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8498
 QY 1432 GluSerGlnAlaGluSerIleProIleValThrProAlaProGluSerThrLeuHis 1451
 Db 8499 CCAAGCCATGCAGAGGG-----CCTCATCTCGCGGAGCAGCGCAT 8540
 QY 1452 ProAspLeuGlnGlyGluLysSerAlaSerGlnArgGluArgSerGluGluAspLys 1471
 Db 8541 CCG-----8543
 QY 1472 ProAspAlaGlyProAspAlaAspGlyLys 1481
 Db 8544 CTTGCTGAGGCGCCAGATCGCCACGCGCGG 8573

RESULT 11
 US-09-724-676A-25032
 ; Sequence 25032, Application US/09724676A
 ; GENERAL INFORMATION:
 ; APPLICANT: CompuGen LTD
 ; TITLE OF INVENTION: Variants of alternative splicing
 ; FILE REFERENCE: 129181.4 CompuGen
 ; CURRENT APPLICATION NUMBER: US/09/724, 676A
 ; CURRENT FILING DATE: 2000-11-28
 ; NUMBER OF SEQ ID NOS: 97222
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 25032
 ; LENGTH: 14335
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-724-676A-25032

Alignment Scores:

Pred. No.: 2.38e-09 Length: 14335
 Score: 418.00 Matches: 334
 Percent Similarity: 35.65% Conservative: 255
 Best Local Similarity: 20.22% Mismatches: 701
 Query Match: 5.18% Indels: 364
 DB: 5 Gaps: 61

US-09-902-432-4 (1-1596) x US-09-724-676A-25032 (1-14335)

Oy SerSerThrluGlulnArgSerProglu-GlnProAlaGlySerAapThrProSerGluLe 24
 Db 4187 AGCAGACAGCGGCAAGAGAGCGCGAGCGCTGCGAGGTGAGAGCGCGCTGAGAAAC 4246
 Oy 24 uValleuSerGlyHsGly-----ProAlaAlaGluAlaSerGlyAlaAlaGlyAapPr 42
 Db 4247 AGCGGAGCTGGCCGAGGCGCAGCGCCAGCAAGGCAAGCGCGGAGCGCGAGAG 4306
 Oy 42 oAlaAspAlaAapProAlaThrluLysleuProGlnLysAenGlyGlnLeuSerSerValAs 62
 Db 4307 AGCTGACAGAGCGCATGCA-----GGAGGAGGTGGTGGCGCGGAGAGCGCG 4354
 Oy 62 nGly-ValAlaGluGlnGlyAspValAlaGlnGluGlnLysAenGlnGlnGlnGlnGln 82
 Db 4355 CGGTGAGCGCGCAGCAGCAGAGCGCAGCATTCAGAGAGAGCTGCGAGCAGCTGCGGACA 4414
 Oy 82 luGluValAlaAspGluAapValAlaGlyGlnArgGluSerGlu----- 95
 Db 4415 GCTCGAGAGCGAGATCCAGGCGCAAGCGCGGCGAGCAGAGCGCGCTGAGCGCAGCGCGC 4474
 Oy 96 --AspValAlaArgGluLysAapArgValAlaGluGlnMetAlaAlaAenSerThrlaValAlaGlu 115
 Db 4475 TCGCATTCAGAGAGAGATCCGCTGGTGGCTGCGCTGCAATTGAGAGCTGAGCGCAGCGCAGC 4534
 Oy 115 spLeThrluLysAapGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 134
 Db 4535 GTCGCGCGGCTGAGGAGGAGCTGCGAGCAGCTGCTGCGAGCGCGGAGAGCGCGAGCGC 4594
 Oy 134 erGluAaAaValAlaGluGlnMetValAlaGlnProAlaGluSerGlnAlaAaAapValAlaGlyP 154
 Db 4595 AAAAGCGAGCGCGCAGCAGAGAGCGCGAGCGCTTGGAGAGCAGGTGCGAGCGAGCGC 4654
 Oy 154 heLysLysValAlaThrluLysPheValAlaGlyPheLysPheThrlaLysLysAap-----L 171
 Db 4655 AGCGTAAAGCGGAGCGGAGGTGAGAGCTGCGCTGCGCTGAAAGCGCGAGCGCGAGCGCG 4714
 Oy 171 yAAsGluLysSerAapThrlaValAlaGlnLeuThrlaLysLysAapGluGlnGluGlu 191
 Db 4715 CGCGGAGAGAGAGCGCGCGCTGCGAGAGCTGCGAGAGCTGCGAGCGCGAGAGAG 4774
 Oy 191 laGlu-----AlaSerValAlaGlyAlaGlyAspHisGlnGlnProSerValG 206
 Db 4775 CGAGAGCGCGCTGCGGAGCGCGAGCTGAGAGCGCGGAGCGAGTAAAGTGGGCTGCTG 4834
 Oy 206 luThr-----AlaValAlaGlyGlnSerAlaSerLysGlnSerGluLeuLysGlnS 222
 Db 4835 AGACGCGCGAGCGCAGCTGAGAGCGGAGCTGCGAGCAAGCGCGCTCTTGGCGGACA 4894
 Oy 222 eThrluLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 241
 Db 4895 AGACGCGCAGCTGAGAGCGCTGCTGCGAGAGAGACACAGTGGCTGGGAGAGCTGCGG 4954
 Oy 241 euGlnAlaGlu-----SerAspGlnAlaAlaGluGlnGluAlaLysAapGluGlyG 258
 Db 4955 AGAGAGCTGAGCGGCGGAGCAGCAGCAGCGAGCGCGAGCGCGCGCGCGAGAGAGCGAG 5014
 Oy 258 luGluLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 278
 Db 5015 AGCGGAGCTGAGCGCTGCGAGCTTCAAGCGCAAGCGAGCGCTTCAAGCGCTGCGGCTGAG 5074
 Oy 278 erGluThrluSerSerPheLysPhePheThrluSerGlyTrpAlaGlyTrpArgLysL 298
 Db 5075 CGAGGAGGTGGCGCAGCAGAGAGCTG-----GCCG 5107

Oy 298 yThrluSerPheLysSerLysGlnLysAapLeuGluThrlaGluLysArgLysGln 318
 Db 5108 AGCGCGAGCTGAGAGAGAGAGAGAGAGAGCGGAGCGCGAGCGCGGCGCGGAGAG 5167
 Oy 318 lngluAlaGluLysValAapGluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 338
 Db 5168 CGAGAGAGAGAGCGCTGCGAG 5224
 Oy 338 lngluProAlaGluAapThrlaValAlaGlnLysSerAlaAapThrluLysValAlaGlu 358
 Db 5225 GCGAGCTGGCGAG 5284
 Oy 358 euProLeuGluAapGlnValAlaLysleuGlnLysAap-----SerGlnGlnLysCysAlaP 377
 Db 5285 TCGGCGCGAG 5344
 Oy 377 roLeu-----AlaThrluValPheAapGlu----- 385
 Db 5345 GGTGACAGCTGAGAGCGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5404
 Oy 386 --LysMetGluAlaHisGlnGlnValAla-----AlaGluValn 398
 Db 5405 CCAAGTTCGAGCGAGATGAG 5464
 Oy 398 lAvalSerThrlaGluLysThrluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 418
 Db 5465 CGGCTTCACAGCAG 5512
 Oy 418 lYValAlaValAlaGluLysThrluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 431
 Db 5513 GCGGTCGCGAG 5572
 Oy 432 -----LysLeuAlaGluProGlnGlnValAlaProGlnGlnAlaGluProAlaGlnG 448
 Db 5573 AGCGGAG 5629
 Oy 448 luLeuMetLysSerArgGluMetCysValSerGlyGlyAspHisThrluLeuThrlaAapL 468
 Db 5630 GGGTCTTGGAG 5674
 Oy 468 euSerProGluGlnLysThrluProLysHisProGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 488
 Db 5675 TCAAG 5734
 Oy 488 eLysSerSerGlnGlu-----ArgLysValAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 503
 Db 5735 GCGTGGAG 5794
 Oy 503 yLysLeuPheSerSer-----SerGlyLeuLysLysLeuSerGly-----LysL 518
 Db 5795 AGGCTGACATCGAG 5854
 Oy 518 ySGlnLysGlyLysArgLysGly-----GlyLysAapGlnGluP 531
 Db 5855 GCGAG 5913
 Oy 531 roGlyGluLysLysHisLysHisThrluSerProGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 549
 Db 5914 CTGGC-----GCTGAAGGCGAGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5961
 Oy 549 ySGlyGlnSerSerLysSerProGlnGlu-ProGlnGluThr-----ThrcysLeuGln 567
 Db 5962 CTGAGAGCTGAG 6021
 Oy 568 LysGlyProLeuGluAlaProGlnAap----- 576
 Db 6022 CAGGCGAGAGCTGAG 6081
 Oy 577 GlyAlaValAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 589
 Db 6082 CGTGAAGCTGAG 6141
 Oy 590 LysArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 609

Db 6142 CGAAGCGCGCTGGAGGAAGTCGACGGCTGAAACCCAACTGGAGGAGCGCGCGC 6201
 Qy ValArgArgProSerGluSerAspLysGluGluLeuLeuLysValLysSerAlaThr 629
 Db 6202 CTGCGGAGCGAGCGAGCAGGAGTCGGCGCGCAGCTGCGAGTGGCCAGGAGCGGCC 6261
 Qy LeuSerSerThrAspSerThrValSerGluMetGlnAspGluValLysThrValGlyGlu 649
 Db 6262 CAGAAGCGCGTG-----CAGGCGAAGAGAACGACACGCGCTTCGCGGTG 6306
 Qy GluGlnLysProGluGluProLysArgValAspThrSerValSerTrpGluAlaLeu 669
 Db 6307 CACAGAGGAGCAGGAGCTACACAGCCCTGCAGCAGGAGCAGAGC-----GTGCTG 6360
 Qy IleCysValGlySerSerLysLysArgAlaArgLysAlaSerSerSerAspGluGly 689
 Db 6361 GACCAGTCGCGCGAGCGGAGCGCGCCGCGCGCGCTGAGGAGCGGAGGAGGCC 6420
 Qy 690 -----GlyProArgThrLeuGlyLysSerHisArgAlaGluGluAlaSerLys 706
 Db 6421 CGGGTCAGCGGAGCGTCGAGCGCGCGCAGCGCGCGCGAGTGGGAAGAGCGCGAGCGG 6480
 Qy 707 AspLysGluAlaGlyThrAspAlaValProAlaSerThrGlnGluGlnAspGlnAlaGln 726
 Db 6481 CTGAAGCAGTCG-----GCAGAGAGCAGGACGACGCGCGCGCTCAG 6522
 Qy 727 GlySerSerProGluProAlaGlySerProSerGluGlyGluGlyValSerThrTrp 746
 Db 6523 GCACAGCGCGCTGCAGAG----- 6540
 Qy 747 GluSerPheLysArgLeuValThrProArgLysLysSerLysSerLysLeuGluGluLys 766
 Db 6541 -----AAGCTGCGCAAGGAGGCGCGAGCAAGAGCGC 6570
 Qy 767 AlaGluAspSerSerValGluGlnLeuSerThrGluLeuGluProSerArgGluGluSer 786
 Db 6571 GCCTGCGCGGCACAGCGCGAGCAGCGCGCTGCGCGAGAGCAGCGAGCTGACGCGGAG 6630
 Qy 787 TrpValSerIleLysLysPheIleProGlyArgArgLysLysArgAlaAspGlyLysGln 806
 Db 6631 ATGGAGNAGCATAGAAATTCGCGCAGCAGACGCTGCGCGAGAGCGCGCAGGTGGAGCAG 6690
 Qy 807 GluGlnAlaThrValGluAspSerGlyProValGluLeuLeuAspGluAspProAsnVal 826
 Db 6691 GAGCTGACACACTCTCGG-----CTGCAGCTGAGGAGAGCCGAC----- 6729
 Qy 827 ProAlaValValProLeuSerGluTyrAsnAlaValGluArgGlu-----Lys 842
 Db 6730 -----CACCAGAAGAACCTGCTGGACGAGGAGCTGCAGCGCGCTGAAG 6771
 Qy 843 MetGlu-----AlaGlnGlyAsnThrGluLeuProGlnLeuLeuGlyAlaVal 858
 Db 6772 GCGGAGCCAGGAGGCGCGCAGCGCAGCGCAGCGAGGAGGAGCTCTCTCGGTG 6831
 Qy 859 TyrValSer---GluGluLeuSerLysThrLeuValHisThrValSerValAlaValIle 877
 Db 6832 CGCGTCAGATGGAGGAGCTGAGCAAG----- 6858
 Qy 878 AspGlyThrArgAlaValThrSerValGluGluArgSerProSerTrpIleSerAlaSer 897
 Db 6859 -----CTCAGGCGCATCATGAGGCTGAGAACCGCGCACTCATCTTTCGCTGACAGGAC 6912
 Qy 898 ValThrGluProLeuGluHisThrAlaGlyGluAlaMetProProVal---GluGluVal 916
 Db 6913 AATACGAGCGCTTCCTGCGAGGAGGAGGCTGAGAAATGATGAAGAGGTGGCGGAGGAGGCC 6972
 Qy 917 ThrGluLysAspIleAlaGluGluThrProValLeuThrGln-----Thr 932
 Db 6973 GCGCGGTGAGTGGCGGCCAAGAGGCTGCGCGACTGCGGAGCTGGCAGAGGAGGAC 7032
 Qy 933 LeuProGluGlyLysAspAlaHisAspAspMetValThrSerGluValAspPheThrSer 952

Db 7033 CTGGCAGCAGCGCGCGCTTGGCAGAGAAGATGCTCAAGCAGAGATGTCAGGCGGTGCAG 7092
 Qy GluAlaValThrAlaThrGluThrSerGluAlaLeuArgThrGluGluValThrGluAla 972
 Db 7093 GAGGCCACCGCAGCTCAAGGCTGAGGCGGAACCTGCTCAGCAGCAGAAGAGGAGCTTGCAG 7152
 Qy SerGlyAlaGluGluThrThrAspMetValSerAlaValSerGlnLeuThrAspSerPro 992
 Db 7153 GAGCAGCGCGCGCGCTGCAGGAG----- 7176
 Qy AspThrThrGluGluAlaThrProValGlnGluValGluSerGlyValLeuAspThrGlu 1012
 Db 7177 GACAAGCAGCAGATGCGCAGCAGCTGCGGAGGAGCAGCAGCGGCTTCACCGCAGCGCTG 7236
 Qy 1013 GluGluGluArgGlnThrGlnAlaIleLeuGlnAlaValAlaAspLysValLys----- 1030
 Db 7237 GAGGCCAGCGCGCAGCGCAGCTGGAGATGAGCGCTGAGCTGAGCGCTCAAGCTCAAGCTGCGT 7296
 Qy 1031 -----GluGluSerGlnValProAlaThrGlnThrValGlnArgThrGlySer 1046
 Db 7297 GTGGCCGAGATGAGCCGAGCCCGCGCTGAGGAGGAGCGCCGAGCGCTTCCGGAAG 7356
 Qy 1047 LysAlaLeuGluLysValGluGluValGluAspSerGluValLeuAlaSerGluLys 1066
 Db 7357 CAGCGCGAGGAGATCGCTGAGAAGCTGCCCGCAGCGAG-----CTGCCACCCAGGAG 7410
 Qy 1067 GluLysAspValMetProLysGlyProValGlnGluAlaGlyAlaGluHisLeuAlaGln 1086
 Db 7411 AAGGTGACCTGTGTGCGAGACACTGGAGATCCAGCAGCAGCAGAGTACCATGATCCGAG 7470
 Qy 1087 GlySerGluThrGlyGlnAlaThrProGluSerLeuGluValProGluValThrAlaAsp 1106
 Db 7471 -----CGCTGCGGAGGAGCGCATCGCTGAG 7494
 Qy 1107 ValAspHis-----ValAlaThrCysGlnValIleLysLeuGlnGlnLeu---MetGlu 1123
 Db 7495 CTGGAGCGTGAGAGGAGAGAGTCCAAACAGAGGAGCGCAAACTGCTGAGCTCAAGTCTGAG 7554
 Qy 1124 GlnAlaValAlaProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThr 1143
 Db 7555 GAGATCAGCAGCGTGCGCAGCAGGAGCAGCTGCTGAGGAGCAGCGCCCTGCAGCAAGC 7614
 Qy 1144 ProLeuAlaAspSerAspThrAlaAspGlyThrGlnGlnAspGluThrIleAspSerGln 1163
 Db 7615 TTCTCTCTGAAAAGACAGCCTG-----CTACAGCGGAGCGCTTCATCGAGCAGGAG 7668
 Qy 1164 AspSerLysAlaThrAlaAlaValArgGlnSerGlnValThrGluGluAlaAlaThr 1183
 Db 7669 AAGGCCAAGCTG-----GAGCAGCTCTTCCAGGAGCAGGTTGCCCAAG 7710
 Qy 1184 AlaGlnLysGluGluProSerThrLeuProAsnAsnValProAlaGlnGluHisGly 1203
 Db 7711 GCACAGCAG-----CTGCGTGAGGAG----- 7731
 Qy 1204 GluGluProGlyArgAspValLeuGluProThrGlnGlnGluThrAlaAlaVal 1223
 Db 7732 CAGCAGCGCGCAGCAGCAGATGAGCAGGAAACCGCAGCGCGCTGCTGGCCAGCATGGAG 7791
 Qy 1224 ProValLeuAlaLysThrGluValGlyGlnGluGly-----GluVal 1237
 Db 7792 GAGCGCGCGCGCGCAGCATGAGCGCGAGGAGGCGCTGCGCGCAGCAGGAGGAGCTG 7851
 Qy 1238 AspTrpLeuAspGlyGluLysValLysGluGluGlnGluValPheValHisSerGlyPro 1257
 Db 7852 CAGCAGCTGAGCAGCAGCGG-----CGGACGAGGAGAGCTGCTGCTGAGGAGAACCCAG 7908
 Qy 1258 AsnSerGlnLysAlaAlaAspValThrTyrAspSerGluValMetGlyValAlaGlyCys 1277
 Db 7909 AGSCTGCTGAGCAGCTGCGAGCTCTCTGAGGAGCAGCAGCGCGCGCTGGCGCACTCA 7968
 Qy 1278 GlnGluLysGluSerThrGluValGlnSerLeuSer----- 1289
 Db 7969 GAGGAGTCACTGCTCGAGGTGGTGGCCAAAGACCCCTGCCCAACCTGCGCGGATGCA 8028

Db 3261 ATACTCAGAAAAATGTGTGACTTCTTAAGTCCCAAGAAATAGAAATTTATTTTAAACAGAC 3320
 QY 363 nValGlyAspLeuGluAla-----SerSerGluGluLysCysAl 376
 Db 3321 AATATCTTGAGAAAGAGTGGAACTACAGCATATAAGGAAGATTTGGAAGAAAG----- 3375
 QY 376 aProLeuAlaThrGluValPheAspGluLysMetGluAlaHisGlnGluValValAlaGl 396
 Db 3376 -----CTGGCAGCTGAA-----GAGCAATTCAGGCTCTGTGTCAACA 3413
 QY 396 uValHisValSerThrValGluLysThrGluGluGlnGlnGlyGlyGlyGluAlaGl 416
 Db 3414 GATGAATCAGACCTTCGAAGATAAACAAACCAA----- 3447
 QY 416 uGlyGlyValValGluGlyThrGlyGlySerLeuProGluLysLeuAlaGluPr 436
 Db 3448 -----ATAGATTTTGCTCCAAAGCAAAATCAGTGAAAA 3479
 QY 436 oGlnGluValProGlnGluAlaGluProAlaGluGluLeuMetLysSerArgGluMetCy 456
 Db 3480 CCAAGCAATATCCAGAAG-----TTAATCACAAGTAAACACGGA 3518
 QY 456 sValSerGlyGlyAspHisThrGlnLeuThrAsp-----LeuSerProGluGl 472
 Db 3519 TGCAGGTGATGGGACTCCGTAGCACTTGTAAAGGAAACAGTGTGTATAGTCCACCTTG 3578
 QY 472 uLysThrLeuProLysHisProGluGlyLeuValSerGluValGluMetLeuSerSerGl 492
 Db 3579 TACAGGTAGTAGTGAACACTGGAACCCAGAACTAGAGAAGAGATACTGCCCCCTTGAAAA 3638
 QY 492 nGluArgGlyLeuValGlnGlySerProLeuLysLysLysLeuPheSerSerSerGlyLeuLy 512
 Db 3639 AGAAAAAGGAGCACTTCAAAGAAGCTACAGGAAGCCTTAACCTCCCGCAAGGCAATCT 3698
 QY 512 sLysLeuSerGlyLysLysGlnLysGlyLysArgGlyGlyGlyLysAspGluProGl 532
 Db 3699 TAAAAAGGCACAGAGAAAGAACATCTCAG-----GAGGAGCTAAA 3743
 QY 532 yGluTyGlnHisIleHisThrGluSerProGluSerAlaAspGluGlnLysGlyGluse 552
 Db 3744 GCAACAGAAAGATGACTATAATCGCTTGCAAGAACAGTTTTGATGAGCAAGCAAGGAAAA 3803
 QY 552 r-----SerAl 554
 Db 3804 TGAGAAATATTGGAGACCAGCTAAGGCAACTCCAGATTCAAGTAAGGGAATCCATAGACGG 3863
 QY 554 aSerSerProGluGluProGluGluThrThrCysLeuGluLysGlyProLeuGluAlaPr 574
 Db 3864 AAAACTCCCAGCAGCAGCAGCAGGAGTGTCTTCCACCTCCAGGTTTAGAAGAAC 3923
 QY 574 oGlnAspGlyGluAlaGluGlyThrThr----- 594
 Db 3924 TTTATTCAAAGCCACAGAACAGCATCACACTCAACCTGTTTTAGAGTCCAACTTGTGCC 3983
 QY 585 -----SerAspGlyGluLysLysArgGluGlyIleThrProTrpAlase 599
 Db 3984 AGACTGCGCTTCTCATCTCAAGATGCGAGTGTCTGCGAGCGGAACCTCTCTGCCCCA 4043
 QY 599 rPheLysLysMetValThrProLysLysArgValArgArgProSerGluSerAspLysGl 619
 Db 4044 GATTAAAGCCAGCTG-----AAGGA 4064
 QY 619 uGluGluLeuGluLysValLys---SerAlaThrLeuSerSerThrAspSerThrValse 638
 Db 4065 AATAGAGCTGAGAAAGTAGAGTTAGAAATTTAGAAAGTTAGTTCTACAACAGTACGCTTAC 4124
 QY 638 rGluMetGlnAspGluValLysThrValGlyGluGluGlnLysProGluGluProLysAr 658
 Db 4125 TAAAAAATCAGAAAGAGGTATTTTCAGTTTACAAAGCAGCATATAATAAACAGGTTTAGAAAT 4184
 QY 658 gArgValAspThrSerValSerTrpGluAlaLeuLysCysValGlySerSerLysLysAr 678
 Db 4185 CGAGAGTCTTAAGACAGTATCCATGAAGCTGAAGTCCATGCGCAAGAGCCTGCAGCAGAA 4244

QY 678 gAlaArgLysAlaSerSerSerAspAspGluGlyGlyProArgThrLeuGlyGlyAspSe 698
 Db 4245 ATTT-----CAAAAGCAGCCAACTACAAATTCGTGGCCTAGA 4280
 QY 698 rHisArgAlaGluGluAlaSerLysAspLysGluAlaGlyThrAspAlaValProAlase 718
 Db 4281 ACATCTAAGAGAAATTCGAACCTTAACCTGGATGAACCTG-----CAAAACTCAATAAG 4331
 QY 718 rThrGlnGluGlnAspGlnAlaGlnGlySer-----SerSerProGluLuproAlaGl 735
 Db 4332 CAAAAAGGAGAGAGAGCTTAGCTTCTCGACAACTTAGTGAGAAAGACAGCTCT 4391
 QY 735 ySerProSerGluGlyGluGlyValSerThrTrpGluSerPheLysArgLeuValThrPr 755
 Db 4392 CACTTAAATACAGACAGAGATAATAGAACAAAGAGATTTAATTAAGGCTCTGCATACA-- 4449
 QY 755 oArgLysLysSerLysSerLysLeuGluLysAlaGluAspSerSerValGluGlnLe 775
 Db 4450 -----CAGCTAGAAATGCAAGCCAAAGAGCATGATGAGAGATTAAGCAGCT 4496
 QY 775 uSerThrGluIle-----GluProSerArgGluGluSerTrpVa 788
 Db 4497 ACAGGTGGAACCTTGTGAATGAACCAAAACCCAGAGAGATTGGAGAGAAAGT----- 4551
 QY 788 lSerIleLysLysPheIleProGlyArgLysLysArgAlaAspGlyLysGlnGlu 808
 Db 4552 -----AGACAAAGCAACAAATACAAAGGAAACTGCAAGC 4586
 QY 808 nAlaThrValGluAspSerGlyProValGluLeuAsnGluAspAspProAsnValProAl 828
 Db 4587 TGCCCTTATTTCCCGAAAA-----GAAGCACTTAAAGAAACAAAGTCTCCAAGA 4637
 QY 828 aValValProLeuSerGluTyAsnAlaValGluArg----- 840
 Db 4638 GGAATGTCTTTGGCCAGA---GGTACCATTGAACGCTCTCACCAGTCTCTGCGCAGATGT 4694
 QY 841 -GluLysMetGluAlaGlnGlyAsnThrGluLeuProGlnLeuLeuLeuAlaValTyVa 860
 Db 4695 GGAAAGCCCAAGTTCTGCTCAAAATAAGAAAGAAAGATACGGTCTTAGGAAGGTTAGCTCT 4754
 QY 860 lSerGluGluLeuSerLysThrLeuValHisThrValSerValAlaValIleAspGlyTh 880
 Db 4755 TCTTCAAGAAAGAAAGACAAACTCTATTACAGAAATGGACAGTCTTTATGGAAATCA 4814
 QY 880 rArgAlaValThrSerValGluGluArgSerProSerTrpIleSerAlaSerValThrGl 900
 Db 4815 GAGTCTCAGCAGCTCTGTGAA-----AGTCTAAAACTAGCTCTAGAGGGTCTTACTGA 4868
 QY 900 uProLeuGluHisThrAlaGlyGluAlaMetProProValGluGluValThrGluLysAs 920
 Db 4869 AGACAAAGGAAAGTTAGTGAAGGAAATTT-----GAATCTTTGAAATCTTTC 4913
 QY 920 pIleAlaGluGluThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHi 940
 Db 4914 TAAGATTGCAAGAAAGTACT-----GAGTGGCAAGAGAAACA 4949
 QY 940 sAspAsp-----MetValThrSerGluValAspPheThrSerGl 953
 Db 4950 CAAGGAGCTTACAAAAGAGATGAAATTTCTTCTGCGAGTCTTATGAGAAATGTAGTAATGA 5009
 QY 953 uAlaValThrAlaThrGluThrSerGluAlaLeuArgThrGluGlu----- 968
 Db 5010 AGCAGAAAGGATTACGATGTGTGGAAGCTGTGAGGCAAGAGAAACAAGAACTGTATGG 5069
 QY 969 -----ValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSerAlaVa 985
 Db 5070 CAAGTTAAGACACAGAGGCAACAAAGAGGAGAGACAGAAAAAGCAGCTGCAGGAAGCTGA 5129
 QY 985 lSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGluValGl 1005
 Db 5130 GCAAGAAATGGAGGAAATGAAAGAAAGATGAGAAAGTTT-----GCTAAATCTTAAACA 5183

```

Qy 1005 userglvalleuasprthrglglugluargrglnthrglnalaileleuglnalava 1025
Db 5184 GCAGAAAATCTTACGCTGAGAGAAAGATGACCGCTTACGAGGAGCTGCACCTGC 5243
Qy 1025 lalaspysvalysglugluserglvalproalathrglnthvalargthrgl 1045
Db 5244 AGAGATACAGCTTAAAGAG-----TGTATGAGAAACCTTCTTCTCCAA 5288
Qy 1045 yserlysalaaleuglulysvalglugluvalglugluasprserglvalleualasergl 1065
Db 5289 TCCCAGCATGAGAGAAAGACTTGAAGGCTCAAAATGAGATGAAACCTTCTTCAAGAA 5348
Qy 1065 u-----lysglulysaspyalmetprolysglprovalglugluagl 1080
Db 5349 GTTTCAGTCTTAAATGCTCTGAGAAAGACTCTTAAAGTAA-----GAGGTTCAAGAT----- 5400
Qy 1080 yalaglulhisleualaglgluserglulthrglglinalathrgluserleugluval 1100
Db 5401 -TTAAAGCATGAGATGAGATGATATCTTAAACAAGCTTAACTGAGGCCACGAGAA 5459
Qy 1100 lprogluvalthralsaspyalaspisvalalathrcysgluval-----lley 1117
Db 5460 ACATGATTAACCAAGCAATGCTCTGAGAGGAGACACAGTCTTACGAGTACGAGACTGA 5519
Qy 1117 sleuglulnleumetgluglnalavalalaprogluserserglulthleuthraspe 1137
Db 5520 AGAGCAAGCTCTCTGATGATGAGCAAGACCTTACATGTTCAAGATCGGTTCCATCAC 5579
Qy 1137 rglulthrasnglyserthproleualasprseraspthralsaspylthrcinglnas 1157
Db 5580 GAAGAGT-----GCCACCCTGCTGTAAGTAAAGATTTCAAGCTCAGATGAAATTAA 5633
Qy 1157 p-----glulthrlleasprserglulnaspserlysalaathralalavalarglnse 1174
Db 5634 TAACATACCTACAGAGATTCATGATGATGATGATGATGATGATGATGATGATGATGAT 5693
Qy 1174 rgluvalthrglugluvalalathralaglnysglugluarproserthleuProas 1194
Db 5694 GCAGAAAACCAAGAAATTTAGCCAGACTTTAGAAATGAGAA-----AATACTTACTGAG 5750
Qy 1194 nasnvalproalaglgluglnlhisglugluarproglulargaspvalleugluProh 1214
Db 5751 TCAGATATCAACAAGAT-----GATGAA-----CTAATAATGCT 5786
Qy 1214 rglnglulnleuthr-----AlaAlaAlaValProva 1225
Db 5787 TCAGAGAGAGATTAACCAAAATGAACTGTTAATACGAAATCAAGAAAGAACTCTCCAG 5846
Qy 1225 lleualalysrthrglulvalgluglulgluvalasprtleuasprglulysva 1245
Db 5847 AGTTACCAAACTTAAGAGACACAGCAAGAAAGAAAGATGATTTGAGAGAGGCTTAT 5906
Qy 1245 llysglulnleugluvalphevalhisserglproasnsrserglulysalalaspva 1265
Db 5907 GAATCAATTAGCAAGACTT-----AATGAAACATTGGGAATTAATCTGATGATGT 5957
Qy 1265 lthr-----TyrAsprserglulvalmetglivalalaglcy 1277
Db 5958 TACAGATGCCCAATTAATAATGAGCTATTTGAAATCTGAATATGAAAGAACTTAAAGTG 6017
Qy 1277 sglnglulysgluserthrglulvalgluserleuglulglulmetglulth 1297
Db 6018 TGTGAGTGAATTGAGAGAAAGAAAGCAAGCAAGTATGCAAGAAATTAAGGTGGAATC 6077
Qy 1297 rlaspyalglulysglulysarglulthrsprroglulnvalserglulglulglul 1317
Db 6078 AGAAAAACCAAGAAATATTGAGAAATATACAAAGTCTCAGAAAGAAACCCGGAATTA 6137
Qy 1317 nglulthralalaproglulhisglulglulthrryrglulysprovalleuthleuaspm 1337
Db 6138 A-----AGCCATGCAAGAAAGAACTTCAGAACTGTTAA 6170
Qy 1337 rProserSerglularglulysalaleugluserleuglulysrProserleuProas 1357

```

```

Db 6171 AGAAAAACCAAGAAATTAAGCACTT-----CAGAAAGACTGCATCAGTA 6218
Qy 1357 polnaspysalaglcyysilegluvalgluvalgluserleuasprthrrvalthrgl 1377
Db 6219 TCAGAGAAAATTAGTGTCTGAGAGAACTTTAAAGCTTGAATTTGTTCAAACTGA 6278
Qy 1377 nthralsglulvalglulysvalleuglulthvalvalleuglulthrglulys 1397
Db 6279 ATCTCAAAAAGATTTGGAATTAACCAAGAAAT-----CTGCTCAAGCAATTGAAACA 6332
Qy 1397 rProglulCysValglulhis-----leuupProalglulys 1411
Db 6333 CCGCAAAAAGCAACAGCAAGATTAAGTACTTCAAACTCTGCTAGATGACACTCAAG 6392
Qy 1411 rserlathrglglulhisrprthrleuglnhisalaglulspthrvalProleuglyPr 1431
Db 6393 TGAAGCAGCA-----AGGCTCTACAGACCAATCTCAAGTTGAAAA 6434
Qy 1431 oguserglulala-----gluserlleproillevalthrproalaprogluserth 1449
Db 6435 GGAACCTCAGTCAATTAAGATCAGTTAAAGCCAGATGAAACAAAGATGA----- 6489
Qy 1449 rleuHisProaspleuglulglulhisserlaserglulnarglulargserglulglul 1469
Db 6490 -----GATCTTGAGCGAGACTGAAACAGCGAGAGAAAGCACTGAAAGAGAA 6539
Qy 1469 uaspylserProasplaglulProasplalaspglulysgluserthralsaleglulysva 1489
Db 6540 GAAGAAATGCAAGAGAAAGTGAATGCTTTGGCGAGAGAAAAAGTCCACTGGAAGAC 6599
Qy 1489 lleu-----lyalaglulproglulileuglulglulys 1501
Db 6600 AATTGAGAGATTCAGGTTACTTTGAAACAGAAAGCAAGAAAGTTGACAACTTCAAG 6659
Qy 1501 rlyserasnlyslleval 1507
Db 6660 AAACCTTGACAGTACTGTG 6678

RESULT 13
US-09-724-676A-7257
Sequence 7257, Application US/09724676A
GENERAL INFORMATION:
/ APPLICANT: Compugen LTD
/ TITLE OF INVENTION: Variants of alternative splicing
/ FILE REFERENCE: 129181.4 Compugen
/ CURRENT APPLICATION NUMBER: US/09/724,676A
/ CURRENT FILING DATE: 2000-11-28
/ NUMBER OF SEQ ID NOS: 97222
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 7257
/ LENGTH: 8161
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-724-676A-7257

Alignment Scores:
Pred. No.: 1,98e-09 Length: 8161
Score: 415.00 Matches: 316
Percent Similarity: 38.17% Conservative: 305
Best Local Similarity: 19.42% Mismatches: 669
Query Match: 5.14% Indels: 337
DB: 5 Gaps: 60

US-09-902-432-4 (1-1596) x US-09-724-676A-7257 (1-8161)
Qy 26 leuserglulys-----HisglulProalalaglulalaserglulalala----- 39
Db 2376 CTTCAAGGTGAAGAACTTACATGATTAACAGAAATTTGAGGCTCAGTAAAGCAACTGGA 2435
Qy 40 gluasprro-AlaaspralaspProalathrlyserleuglulysaenglulglulys 59
Db 2436 AATGAACCTTGCAAGAGCAAGAAAGCAAGAAAGAACTGATTTATGAAGC---CAAACTGC 2492

```

Qy	59	rSerValAsnGlyValAlaGlnGlnClyAspValHisValGlnGlnGluAsnGlnGluGl	79
Db	2493	CCATGACAACTCCTCAGTGAACAG-----ATCCATAGTCTCAGCATAGAACGCAATC	2546
Qy	79	YGlnGluGluValValAspGluAspValGlyGlnArgGlnSerGluAspValArgGl	99
Db	2547	TAAAGATGCTGAATTT-----GAAGTTTACAGAAATGAAGTGAATGTGCAG--	2595
Qy	99	uLysAspArgValGluGluMetAlaAlaAsnSerThrAlaValGluAspIleThrLysAs	119
Db	2596	-----CTTCAGTTTTCTCAGCAGAGTAGTACCCTGATAAGAAAGCTGCAAAGCCA	2642
Qy	119	pGlyGlnGluGluThrSerGluIleIle-----GluGlnIleProAlaserGluAs	136
Db	2643	GCTGCAAAATAAGAAAGTGAAGTCTTGAGGGGCGCAACGCTGTAAAGGCATATCTCAAG	2702
Qy	136	nAsnValGluGluMetValGlnProAlaGluSerGlnAlaAsnAspValGlyPheLysLy	156
Db	2703	TAAAGTGAAGAACTGCTCCAGGCTTTTCACAGAAGAACTTGAATA-----	2751
Qy	156	sValPheLysPheValGlyPheLysPheThrValLysLysAspLysAsnGluLysSerAs	176
Db	2752	-----ACAAAAATGCAT-----	2763
Qy	176	pThrValGlnLeuLeuThrValLysLysAspGluGlyGluGlyAlaGluAlaSerValGl	196
Db	2764	-----CAGCTCTTACTAGAAAAAGAGATGTGGAAACCTCTCAACAAACCATCGA	2816
Qy	196	yAlaGlyAspHisGlnGluProSerValGluThrAlaValGlyGluSerAlaSerLysGl	216
Db	2817	GGAGAAGGATCACAACTGACAGAAATCAGCTTTAGTATGACTGAGAAATGTTTCAGCT	2876
Qy	216	uSerGluLeuLysGlnSerThrGluLysGlnGluGlyThrLeuLysGlnGluGlnSerSe	236
Db	2877	TAATGAAGAGAACTTTCTCTCTGGGGTTGAAATTAAGACTCTTAAAGAA-----	2925
Qy	236	rThrGluIleProLeuGlnAlaGluSerAspGlnAlaAlaGluGluGluAlaLysAspGl	256
Db	2926	-----CAGCTAAATTTATTCACAGAGCTGAGAGCAAAAAAGAGCAGGTGGGAAGA	2981
Qy	256	uGlyGluGluLysGlnGluGlyProThrLysSerProGluSerProSerProVa	276
Db	2982	TAATGAAGTTTCTCTGGCTTTAAACAAATTTATCATGATGAGCCAGCCAGGACAAAT	3041
Qy	276	lAsnSerGluThrThrSerSerPheLysLysPhePheThrHisGlyTrpAlaGlyTrpAr	296
Db	3042	AAGTAGGAGGAA-----CTTCAGCATGAATTTGACCTTTCTGAA	3080
Qy	296	gLyLysThrSerPheLysLysSerLys-----GluAspAs	308
Db	3081	GAAGAAGAAATGACGACAGAAAGAGAAAGCTCCAGGCAGCTCTTTATTACAGAAAGGAGCT	3140
Qy	308	pLeuGluThrAlaGluLysArgLysGlnGlnAlaGluLysValAspGluGluGluLy	328
Db	3141	TCTGCAAGAGTCAGTAGATTGGAGAGAAATTAACCAACTTGAAGATGAACTAAGAA	3200
Qy	328	sGluLysThrGluProAlaSerGluGlnGluProAlaGluAspThrAspGlnAlaAr	348
Db	3201	AGAAATCCCACTCAGTCAGATGAGAGGGGGAAGTGAAGAGAAGATAAGAAAAAAGAA	3260
Qy	348	gLeuSer-----AlaAspTyrGluLysValGluLeuProLeuGluAspGl	363
Db	3261	ATACTCAGAAAAATGTGTGACTTCTAAGTGGCCAGAAATAGAAATTTATTTTAAACAGAC	3320
Qy	363	nValGlyAspLeuGluAla-----SerSerGluGluLysCysAl	376
Db	3321	AATATCTGAGAAAGAAAGTGGAACTACAGCATATAGGAAGGATTTTGAAGAGAAAG	3375
Qy	376	aProLeuAlaThrGluValPheAspGluLysMetGluAlaHisGlnGluValValAlaGl	396
Db	3376	-----CTGGCAGCTGAA-----GAGCAATTCAGGCTCTGTGTCACAA	3413

Qy	396	uValHisValSerThrValGluLysThrValGluGluGluGlnGlyGlyGlyGluAlaG1	416
Db	3414	GATGAATCAGACCTTCAGAAATAAAACAAACCA	3447
Qy	416	uGlyValValGluGlyThrGlyGluSerLeuProGluLysLeuAlaGluPr	436
Db	3448	-----ATAGATTGTCTCAACAGCAAAATCAGTGAAAA	3479
Qy	436	oGlnGluValProGlnGluAlaGluProAlaGluLeuMetLysSerArgGluMetCy	456
Db	3480	CCAAGCAATTATCCAGAAG-----TTAATCACAAGTAACACGGA	3518
Qy	456	sValSerGlyGlyAspHisThrGlnLeuThrAsp-----LeuSerProGluG1	472
Db	3519	TGCAAGTGCATGGGACCTCGTAGACATCTGTAAAGAAACAGCTGTGTATAAGTCCACCTTG	3578
Qy	472	uLysThrLeuProLysHisProGluGlyLeuValSerGluValGluMetLeuSerSerG1	492
Db	3579	TACAGGTAGTAGTAACACTCGGAACCGAAGACTAGAAAGAAATATCTGGCCCTTGAAAA	3638
Qy	492	nGluArgIleLeuValGlnGlySerProLeuLysLysLeuPheSerSerGlyLeuLys	512
Db	3639	AGAAAAGGACCACTTCAAAAGAGCTACAGGAAGCTTAACTCCCGCAAGCAATTCT	3698
Qy	512	sLysLeuSerGlyLysLysGlnLysGlyLysArgGlyGlyGlyGlyGlyGluProG1	532
Db	3699	TAAAAAGGCAGAGAGAAAGAAACATCTCAGG-----GAGGAGCTAAA	3743
Qy	532	yGluTyrGlnHisIleHisThrGluSerProGluSerAlaAspGluGlnLysGlyGluSe	552
Db	3744	GCAACAGAAGAGTAGCTATAATCGCTTGCAAGAACAGTTGTATGAGCAAGCAAGGAAAA	3803
Qy	552	r-----SerAl	554
Db	3804	TGAGAAATTTGGAGACAGCTAAGCCAACCTCCAGATTCAAGTAGGGAATCCATAGACGG	3863
Qy	554	aSerSerProGluGluProGluGluThrThrCysLeuGluLysGlyProLeuGluAlaPr	574
Db	3864	AAAACTCCCAAGCACAGACAGCAGCAAGAAATCGTGTTCTTCCACTCCAGGTTTGAAGAACC	3923
Qy	574	oGlnAspGlyGluAlaGluGluGlyThrThr-----	584
Db	3924	TTTATTCAAGCCACAGAACACGATCACAACCTGTTTGTAGAGTCCAACTTGTGCCCC	3983
Qy	585	-----SerAspGlyGluLysLysArgGluGlyLeuThrProTrpAlase	599
Db	3984	AGACTGGCCTTCTCATCTCGAGATCGCAGTGCTCTGCAGCGCGGAAGTCTGTGTGCCCA	4043
Qy	599	rPheLysLysMetValThrProLysLysArgValArgArgProSerGluSerAspLysG1	619
Db	4044	GATTAAAGGCCAGCTG-----AAGGA	4064
Qy	619	uGluGluLeuGluLysValLys-----SerAlaThrLeuSerSerThrAspSerThrValse	638
Db	4065	AATAGAGGCTGAGAAAGTAGAGTTAGAAATGAAAGTTAGTTCTACAAACAAGTCGAGCTTAC	4124
Qy	638	xGluMetGlnAspGluValLysThrValGlyGluGluGlnLysProGluGluProLysAr	658
Db	4125	TAAAAAATCMAGAGAGTATTTTCAGTTACAGAGCAGCATATAATAAACAGGGTTTGAANAAT	4184
Qy	658	gArgValAspThrSerValSerTrpGluAlaLeuIleCysValGlySerSerLysLysAr	678
Db	4185	CGAGAGTCTAAAGACAGATATCCATGAAGCTCCATGCCGAAAGCTGCAGCAGAA	4244
Qy	678	gAlaArgLysAlaSerSerSerAspAspGluGlyGlyProArgThrLeuGlyGlyAspSe	698
Db	4245	ATTG-----GAAAGCAGGCCAACTACAAATTTGCTGGCCTAGA	4280
Qy	698	rHisArgAlaGluGluAlaSerLysAspLysAlaGlyThrAspAlaValProAlase	718
Db	4281	ACATCTAAGGAATTCACCACTTAACCTGGATGAACCTG-----CAAAAACTCATAAG	4331
Qy	718	rThrGlnGluGlnAspGlnAlaGlnGlySer-----SerSerProGluProAlaG1	735

Db 4332 CAAAAGGAAGAGCTTACCTTCTTGACAACTTATGAGAGAAAGAACCTCT 4391
 QY 735 ySerProserGluGluGluGluValSerThrTrpGluSerPheLysArgLeuValThrPr 755
 Db 4392 CACTAAATATACAGACAGATTAATAGAACAGAAAGTTTAATTAAGCTCTGCATACA-- 4449
 QY 755 oArgLysLysSerLysSerLysLeuGluGluValGluAspSerSerValGluGluLe 775
 Db 4450 -----CAGCTAGAAATGCAAGCCAAAGACATGATGAGGATTAAGAGCT 4496
 QY 775 userThrGluLe-----GluProSerArgGluGluSerTrpVal 788
 Db 4497 ACAGGTGAACTTGTGAAATGAAAGCAAAACCAAGAGATGAGAGAAAGT----- 4551
 QY 788 lSerLysLysPheLeuProGluArgLysLysArgLysAspLysGluGluGlu 808
 Db 4552 -----AGAGCAAGCAAAATGCAAAAGGAACTGCAAGC 4586
 QY 808 nAlaThrValGluAspSerGlyProValGluLeuGluAspAspProAsnValProAl 828
 Db 4587 TGGCTTATTTCCCGAAA-----GAGCACTAAAGAAAGAAAGAAAGTCTCCAGA 4637
 QY 828 aValValProLeuSerGluTyrAsnAlaValGluArg----- 840
 Db 4638 GGAATGTCTTGTGGCAGA---GCTACATGAACTGCTCACCAGTCTTGGCAGATGT 4694
 QY 841 -GluLysMetGluAlaGlnGlyAsnThrGluLeuProGluLeuLeuGluValAlaValTyrVa 860
 Db 4695 GGAAGCCCAAGTTTCTGCTCAAAATTAAGAAAGAAATACGCTTGTGAAAGTTAGCTCT 4754
 QY 860 lSerGluGluLeuSerLysThrLeuValHisThrValSerValAlaValLysGluGlu 880
 Db 4755 TCTTCAAGAAAGAAAGAGACAAACTCATTAAGAAATGAGACAGCTTATTAAGAAATCA 4814
 QY 880 rArgAlaValThrSerValGluGluGluArgSerProSerTrpLysSerAlaSerValThrGl 900
 Db 4815 GAGCTCAGACAGCTCTGTGAA-----AGTCTAAACTAGCTCTAGAGGCTTACTGA 4868
 QY 900 uProLeuGluHisThrAlaGlyLysLysLysMetProProValGluGluValThrGluLysAs 920
 Db 4869 AGCAAGAGAAAGATTAGTGAAGAAATT-----GAATCTTGAATCTTC 4913
 QY 920 pIleIleAlaGluGluThrProValLeuThrGluThrLeuProGluGlyLysAspAlaHis 940
 Db 4914 TAAGATTGCAAGAAATGACT-----GAGTGGCAAGAGAAACA 4949
 QY 940 sAspAsp-----MetValThrSerGluValLysPheThrSerGly 953
 Db 4950 CAAGAGGCTACAAAAGATGATAATTTCTTGCAGCTCTATGAGAAATGTATGATATGA 5009
 QY 953 uAlaValThrAlaThrGluThrSerGluAlaLeuArgThrGluGlu----- 968
 Db 5010 AGCAGAAAGGATTACAGATGTGTGAGAGCTGTGAGGCAAGAGAAACAAGAACTGATGG 5069
 QY 969 -----ValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSerAlaVal 985
 Db 5070 CAAGTTAAGAGCACAGAGGCAAAAGAAAGAGACAGAAAGACAGTTGAGAGAACTGA 5129
 QY 985 lSerGluLeuThrAspSerProAspThrThrGluGluAlaThrProValGluGluValGl 1005
 Db 5130 GCAGAAATGAGGAATGAAAGAAAGATGAGAAAGTTT-----GCTAAATCTTAAACA 5183
 QY 1005 userGlyValLeuAspThrThrGluGluGluGluValArgGlnThrGluAlaLysLeuGluAlaVal 1025
 Db 5184 GCAGAAATCTTACAGCTGAGAGAGAAAGATGACCGCTTAAGGAGAGAGTGCACCTGC 5243
 QY 1025 lAlaAspLysValLysGluGluSerGluValProAlaThrGlnThrValGlnArgThrGl 1045
 Db 5244 AGGAGATTACAGCTAAAGAG-----TGTATGAAACACCTCTTCTTCCAA 5288
 QY 1045 ySerLysAlaLeuGluLysValGluGluValGluGluAspSerGluValLeuAlaSerGly 1065

Db 5289 TGCACAGATGAAGAGAACTTGAAGGGTCAAAATGAGATATGAACCTTTCTTAAGA 5348
 QY 1065 u-----LysGluLysAspValMetProLysGlyProValGlnGluAlaGly 1080
 Db 5349 GTTTCAGTCTTAAATGTCTGGAAGAAAGACTCTTAAGTAA---GAGGTTCAGAT----- 5400
 QY 1080 yAlaGluLysLeuAlaGlnGlySerGluThrGlyGlnAlaThrProGluSerLysLeuVal 1100
 Db 5401 -TTAAAGCATCGATAGAGAGGTAATGTATCTTAACACACTTAACCTAGAGGCCACCGAGAA 5459
 QY 1100 lProGluValThrAlaAspValLysPheValAlaThrCysGluGlnVal-----Lys 1117
 Db 5460 ACATGATTAACCAAGATGTCTAGAGAGGAAACACAGTCTATACAGAGTGAAGACGTGA 5519
 QY 1117 lLeuGlnGluLeuMetGluGlnAlaValAlaProGluSerSerGluThrLeuThrAspSe 1137
 Db 5520 AGAGCAAGCTCTCTGAGTATGAGCAAGACCTTCAAGTTTCAGAAATGCGTTCCATCAGC 5579
 QY 1137 rGluThrAsnGlySerThrProLeuAlaAspSerAspThrAlaAspGlyThrGlnGlnAs 1157
 Db 5580 GAAGAGT-----GCCAACCTGCTGTAGTAAAGATTTTCACTCAGATGAATTTAA 5633
 QY 1157 p-----GluThrLysAspSerGluAspSerLysAlaThrAlaAlaValArgLys 1174
 Db 5634 TAACCTACCTACAGAGATTTGATCAGCTCAAGAAAGAAATGCTGATTAAGAGAGAGAA 5693
 QY 1174 rGluValThrGluGluGluAlaAlaThrAlaGlnLysGluGluProSerThrLeuProAs 1194
 Db 5694 GCAGAAAGCAAGAAATTTAGCCAGACCTTGAAGAAATGAGAAA---AATACCTTACTAG 5750
 QY 1194 nAsnValProAlaGlnGluGluHisGlyGluGluProGluLysArgAspValLeuGluProTh 1214
 Db 5751 TCACTATCAACAAGAGAT-----GATGAA-----CTAAAGATGCT 5786
 QY 1214 rGlnGlnGluLeuThr-----AlaAlaAlaValProVal 1225
 Db 5787 TCAGAGAGAAAGTAAACAAATGAACCTGTAAATCAGCAAAATCCAAAGAACTCCAG 5846
 QY 1225 lLeuAlaLysThrGluValAlaGlyGlnGluGlyGluValAspTrpLeuAspGlyLysVal 1245
 Db 5847 AGTTACCAAACTAAAGAGAGACAGCAAGAAAGAAAGATGATTGGAAAGAGAGCTTAT 5906
 QY 1245 lLysGluGluGlnGluValPheValHisSerGlyProAsnSerGlnLysAlaAlaAspVal 1265
 Db 5907 GAATCAATTAGCAAGACTT-----AATGGAAGCAATGGAAATTACTGTACAGATGT 5957
 QY 1265 lThr-----TyrAspSerGluValMetGlyValAlaGlyCys 1277
 Db 5958 TACAGATGCCAAATTAATAAATAGACTATTTGGAATCTGAATGAAGAACCTTAATAAGTG 6017
 QY 1277 rGlnGluLysGluSerThrGluValGlnSerLeuSerLeuGluGluGlyGluMetGluTh 1297
 Db 6018 TGTGAGTGAATTGGAAGAAAGAAACACAGCTTACTCAAGAAAGAACTAAAGTGAATC 6077
 QY 1297 rAspValGluLysGluLysArgGluThrLysProGluGlnValSerGluGluGluGluGlu 1317
 Db 6078 AAAAAATACAAAGGAATATTGGAGAAAAATACAGGTGCTCAGAAAGAACCCGGAATTA 6137
 QY 1317 nGluThrAlaAlaProGluHisGluGlyThrTyrGlyLysProValLeuThrLeuAspMe 1337
 Db 6138 A-----AGCATGCAAAAGAACTTCAGAAACTGTAA 6170
 QY 1337 rProSerSerGluArgGlyLysAlaLeuGlySerLeuGlyLysSerProSerLeuProAs 1357
 Db 6171 AGAAAAACAACAAGAAAGTAAAGACACTA-----CAGAAGCACTGCATCAGTA 6218
 QY 1357 rGluAspLysAlaGlyCysLysGluValGlnValGlnSerLeuAspThrThrValThrGl 1377
 Db 6219 TCAGAGAAATTAAGTGTCTTGGAGAGAACTGTTAAGCTTGAAGATTGTTCAAACTGA 6278
 QY 1377 nThrAlaGluAlaValGluLysValLysGluThrValLysSerGluThrGlyGlyLys 1397
 Db 6279 ATCTCAAAAAGATTGGAAATTAACCAAGAAAT-----CTGCTCAAGAGTTGAACA 6332

```

QY 1397 rProGluCysValGlyAlaHis-----LeuLeuProAlaGluLysSe 1411
Db 6333 CCACAAGAGGCAACAGCAGAAATTAGTAGCTTCAAAGTCTCTGTAGATGACACTCAAAG 6392
QY 1411 rSerAlaThrGlyGlyHisrPrThrLeuGlnHisAlaGluAspThrValProLeuGlyPr 1431
Db 6393 TGAAGCAGCA-----AGGTCCTAGCAGACAACTCAAGTTGAAAAA 6434
QY 1431 oGluSerGlnAla-----GluSerIleProIleValThrProAlaProGluSerTh 1449
Db 6435 GGAACCTTCAGTCAAATAAAGAAATCAGTTAAAGCCAGATGAAACAAAGGATGAA----- 6489
QY 1449 rLeuHisProAspLeuGlnGlyGluLeuSerAlaSerGlnArgGluArgSerGluGlu 1469
Db 6490 -----GATCTTGAGCGAAGACTGGAAACAGCGCAGAGAGAACCACTGAAAGAGAA 6539
QY 1469 uAspLysProAspAlaGlyProAspAlaAspGlyLysGluSerThrAlaIleGluLysVa 1489
Db 6540 GAAGAAATATCAAGAGAAACTGGATGCTTGGCAGAGAAAAAGTCCACTTGGAGAGAC 6599
QY 1489 lLeu-----LysAlaGluProGluIleLeuGluLeuGluSe 1501
Db 6600 AATTGGAGAGATTTCAGTTACTTTTGAACAAGAAAGACAAAGAAAGTTTCAGCACTTCAGGA 6659
QY 1501 rLysSerAsnLysIleVal 1507
Db 6660 AAACCTTGGACAGTACTGTG 6678

RESULT 14
US-09-724-676-7240
; Sequence 7240, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7240
; LENGTH: 8189
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676-7240

Alignment Scores:
Pred. No.: 1 98e-09 Length: 8189
Score: 415.00 Matches: 316
Percent Similarity: 38.17% Conservative: 305
Best Local Similarity: 19.42% Mismatches: 669
Query Match: 5.14% Indels: 337
DB: 5 Gaps: 60

US-09-902-432-4 (1-1596) x US-09-724-676-7240 (1-8189)
QY 26 LeuSerGly-----HisGlyProAlaAlaGluAlaSerGlyAlaAla----- 39
Db 2404 CTCTCAGGTGAAGAACTTAGCATGGTAACAGAAATTGAGGGCTCAGTAAAGCACTGGA 2463
QY 40 GlyAspPro-AlaAspAlaAspProAlaThrLysLeuProGlnLysAsnGlyGlnLeuSe 59
Db 2464 AATGAACCTTGCAGAGAGAGAGGAGGAGAAAGACACTTATTATGAAAGC---CAAAGTGC 2520
QY 59 rSerValAsnGlyValAlaGluGlnGlyAspValHisValGlnGluAsnGlnGlu 79
Db 2521 CCATGCAACCTCTGCTACTGAACAG-----ATCCATAGTCTCAGCATAGAAGCCAAATC 2574
QY 79 yGlnGluGluValValAspGluAspValGlnArgGluSerGluAspValArgG 99
Db 2575 TAAAGATGTGAATTT-----GAAGTTTACAGATGAATGATGTGTCGAG-- 2623
QY 99 uLysAspArgValGluGluMetAlaAlaAsnSerThrAlaValGluAspIleThrLysAs 119

```

```

Db 2624 -----CTTCAGTTTTCTGAGCAGAGTAGTACCCTGTAAAGAGCCTCCAAAGCCA 2670
QY 119 pGlyGlnGluGluThrSerGluIleIle-----GluGlnIleProAlaSerGluAs 136
Db 2671 GCTGCAAAATAAGAAAGTGAAGTCTTCAGGGGCGCAGAACGCTGTAAAGCATATCTCAAG 2730
QY 136 nAenValGluGluMetValGlnProAlaGluSerGlnAlaAsnAspValGlyPheLys 156
Db 2731 TAAAGTGAAGAACTGCTCCAGGCTCTTTCACAGAGGAAGAACTTGAATA----- 2779
QY 156 sValPheLysPheValGlyPheLysPheThrValLysLysAspLysAsnGluLysSerAs 176
Db 2780 -----ACAAAATATGGAT----- 2791
QY 176 pThrValGlnLeuLeuThrValLysLysAspGluGlyGluGlyAlaGluAlaSerValG 196
Db 2792 -----CAGCTTCTACTAGAGAAAGAGAGATGTGGAACCCCTCCACAAACCATCCA 2844
QY 196 yAlaGlyAspHisGlnGluProSerValGluThrAlaValAlaGlyGluSerAlaSerLys 216
Db 2845 GGAGAGAGGATCAACAAGTGCACAGAAATCAGCTTTAGTATGACTGAGAAAATGGTTTCAGCT 2904
QY 216 uSerGluLeuLysGlnSerThrGluLysGlnGluGlyThrLeuLysGlnGluGlnSerSe 236
Db 2905 TAATGAAGAGAAAGTTTCTCTGGGTTGAAATTAAGACTCTTAAAGAA----- 2953
QY 236 rThrGluIleProLeuGlnAlaGluSerAspGlnAlaAlaGluGluGluAlaLysAspG 256
Db 2954 -----CAGCTAAATTTATTTATCCAGAGCTGAGGAGCAAAAAGAGCAGCGTGAAGAGA 3009
QY 256 uGlyGluGluLysGlnGluGluProThrLysSerProGluSerProGluSerProSerProVa 276
Db 3010 TAATGAAGTTTCTCTGGCTTAAACAAATTTATGATGAGATGAGATGAGCCAGCAGGACAAAT 3069
QY 276 lAenSerGluThrThrSerSerPheLysLysPhePheThrHisGlyTrpAlaGlyTrpAr 296
Db 3070 AAGTGAAGGAGAA-----CTTCAGCATGAATTTGACCTTCTCGAA 3108
QY 296 gLysLysThrSerPheLysLysSerLys-----GluAspAs 308
Db 3109 GAAAGAAATAGCAGAGAGAAAGAGAAAGCTCCAGGCGCTCTTTTAAACAGAAAGGAGCT 3168
QY 308 pLeuGluThrAlaGluLysArgLysGlnGlnGluAlaGluLysValAspGluGluGlu 328
Db 3169 TCTGCAAGAGCTCAGTAGATTGGAAGAGAAATTAGCCAATTTGAAAGATGAATCTAAGAA 3228
QY 328 sGluLysThrGluProAlaSerGluGlnGluProAlaGluAspThrAspGlnAlaAr 348
Db 3229 AGAAATCCCACTCAGTGAGACTCAGAGGGGAGAGAGTGAAGAGAGATAAAGAAACAAAGA 3288
QY 348 gLeuSer-----AlaAspTyrGluLysValGluLeuProLeuGluAspG 363
Db 3289 ATACTCAGAAATATGTGCTACTTCTAAGTGCAGAAATATAGAAATTTATTTAAACAGAC 3348
QY 363 nValGlyAspLeuGluAla-----SerSerGluGluLysCysAl 376
Db 3349 AATACTGAGAAAGAGTGAAGTGAAGTACATATAGGAAGGATTTTGAAGAAAG----- 3403
QY 376 aProLeuAlaThrGluValPheAspGluLysMetGluAlaHisGlnGluValValaG 396
Db 3404 -----CTGCGAGCTGAA-----GAGCAATTCAGGCTCTGCTCAACA 3441
QY 396 uValHisValSerThrValGluLysThrGluGlnGlnGlyGlyGlyGlyGluAlaG 416
Db 3442 GATGAATCAGACTTTCAGAGATANAACCAACCA----- 3475
QY 416 uGlyGlyValValGluGlyThrGlyGluSerLeuProGluLysLeuAlaGluPr 436
Db 3476 -----ATAGATTGCTCCAGCAGAGAAATCAGTGAAAA 3507
QY 436 oGlnGluValProGlnGluAlaGluProAlaGluLeuMetLysSerArgGluMetCy 456

```


Db 3508 CCAAGCAATTATCCAGAG-----TTATCAACAAGTAACGGA 3546
 QY 456 eValserGlyIAspHisThrGlnLeuThrAsp-----LeuSerProGluI 472
 Db 3547 TCCAAGTGAATGGAGACTCCGTAGCACTGTAAAGAAAGAGTGGTATGATCCACCTTG 3606
 QY 472 uLysThrLeuProLysHisProGluGluIleValSerGluValGluLeuSerGluI 492
 Db 3607 TACAGGTAGTACTGAACTGGAAACCCAGAACCTAGAAAGAAAGATCTGCGCTTGAANA 3666
 QY 492 ngLysArgIleLysValGlnGlySerProLeuLysLysLeuPheSerSerGlyLeuLys 512
 Db 3667 AGAAAAGAGCAACTTCAAAAAGAGCTACAGAAAGCTTAACCTCCGCAAGCAATTCCT 3726
 QY 512 sLysLeuSerGlyLysLysGlnLysGlyLysArgGlyGlyGlyLysAspGluProGlu 532
 Db 3727 TAAAAAGCAGCAGAGAAAGAAAGAACATCTCAAG-----GAGAGCTTAA 3771
 QY 532 yGluTyrgInHisIleHisThrGluSerProGluSerAlaAspGluGlnLysGlyLys 552
 Db 3772 GCACAGAAAGATGACTATATCTTGCAAGAACAGTTGATGACCAAGCAAGGAAA 3831
 QY 552 r-----SerAl 554
 Db 3832 TGAGATATTGGAGACCACTAGCAACTCCAGATTCAAGTAAGGAATCCATAGCG 3891
 QY 554 aSerSerProGluGluProGluGluThrThrCysLeuGluLysGlyProLeuGluAlaPr 574
 Db 3892 AAAATCCCAAGCAGACAGCAGCAAGAAATCGTGTCTTCACCTCCAGGTTAGAAAGCC 3951
 QY 574 oGlnAspGlyGluAlaGluGluGlyThr-----584
 Db 3952 TTTATTCAAAAGCCACAGAACAGATCACTCACTGTTTAGAGTCCAACTTGTCCTCC 4011
 QY 585 -----SerAspGlyGluLysArgGluGlyIleThrProTriPylase 599
 Db 4012 AGACTGCGCTTCATTCATTCGAATGCAAGTGTCTGCGAGGCGGAATCTTGTCCTCA 4071
 QY 599 rPheLysLysMetValThrProLysLysArgValArgArgProSerGluSerAspLysGlu 619
 Db 4072 GATTAGGCCCAAGCTG-----AAGGA 4092
 QY 619 uGluGluLeuGluLysValLys-----SerAlaThrLeuSerSerThrAspSerThrValSe 638
 Db 4093 AATGAGGCTGAGAAAGTGAAGTGAATGAAGTTAGTTAGTTACAAACAAGTGAAGTTTAC 4152
 QY 638 rGluMetGlnAspGluValLysThrValGlyGluGluGlnLysProGluGluProLysAr 658
 Db 4153 TAAAAAATCAGAAAGAGTATTTCAGTTACAGAGCAGATTAATAACAGAGTTTAGAAT 4212
 QY 658 gArgValAspThrSerValSerTriPylaseLysLysValGlySerSerLysLysAr 678
 Db 4213 CCGAGATCTTAAGACAGTATCCATGAGTGAAGTCCAGGAAAGCTTCGACAGAGA 4272
 QY 678 gAlaArgLysAlaSerSerSerAspAspGluGlyGlyProArgThrLeuGlyLysPse 698
 Db 4273 ATTG-----GAAGCAGCAACTCAAAATTCCTGGCTTAGA 4308
 QY 698 rHisArgAlaGluGluAlaSerLysAspLysGluAlaGlyThrAspAlaValProAlaSe 718
 Db 4309 ACATCTTAAGGAATTGCAACCTTAACCTGATGACTG-----CAAAAATCTTAAG 4359
 QY 718 rThrGlnGluGlnAspGlnAlaGlnLysSer-----SerSerProGluProAlaGlu 735
 Db 4360 CAAAAGGAAGAGAGCTTAGTACTTCTGTGACAACCTTAGAGAGAAAGAACACACTCT 4419
 QY 735 ySerProSerGluGlyGlyGlyValSerThrTriPylasePheLysArgLeuValThrPr 755
 Db 4420 CACTAAATAATACAGACAGATTAATAGAACAAAGAAATTAATTAAAGCTCTTGATACA-- 4477
 QY 755 oArgLysLysSerLysSerLysLeuGluGluLysAlaGluAspSerSerValGluGlnLe 775
 Db 4478 -----CAGCTAGAAATGCAAGCCAAAGACATGATGAGAGATTAAGCAGCT 4524

QY 775 userThrGluLe-----GluProSerArgGluGluSerTrpVa 788
 Db 4525 ACAGGTGAACCTTGTGAATGAAGCAAAAACCAAGAAAGATGGAGAGAAAGT----- 4579
 QY 788 lSerIleLysLysPheIleProGluArgArgLysLysArgAlaAspGlyLysGlnGluI 808
 Db 4580 -----AGGCAAGCAACAATTAACAAAGAAAGAACTGCAGC 4614
 QY 808 nAlaThrValGluAspSerGlyProValGluIleAsnGluLysPheProAsnValProAl 828
 Db 4615 TGCCCTTATTTCCGAAAA-----GAAGCACTAAAGAAAGAAAGAAAGCTCCAGAA 4665
 QY 828 aValValProLeuSerGluTyrgAsnAlaValGluArg-----840
 Db 4666 GGAATGTCTTGGCCACA--GGTACCATTAAGCTCTCAACCAAGTCTGGCAGATGT 4722
 QY 841 -GluLysMetGluAlaGlnGlyAsnThrGluLeuProGlnLeuGluGlyAlaValTyrgVa 860
 Db 4723 GAAAAGCCAAAGTTTGTCTGCTCAAAATTAAGAAAGATACGGTCTTAGAGAGTTAGCTCT 4782
 QY 860 lSerGluGluLeuSerLysThrLeuValHisThrValSerValAlaValIleAspGlyTh 880
 Db 4783 TCTTCAAGAAAGAAAGACAAACTCATTAACAAGATGACAGGTCTTATTGAAATCA 4842
 QY 880 rArgAlaValThrSerValGluGluArgSerProSerTriPylaseSerAlaThrGluI 900
 Db 4843 GAGTCTCAGCACTCTCTGTGA-----AGCTTAAACTAGCTCTAGAGGCTTTCAGCA 4896
 QY 900 uProLeuGlnHisThrAlaGlyGluAlaMetProProValGluGluValThrGluLysAs 920
 Db 4897 AGACAAGAAAGAAAGTGAAGAAAT-----GAATCTTGAATCTTC 4941
 QY 920 pIleIleAlaGluGluThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHi 940
 Db 4942 TAAATGTGCAAGAAAGTACT-----GAGTGCAAGAAAGAAACA 4977
 QY 940 eAspAsp-----MetValThrSerGluValAspPheThrSerGlu 953
 Db 4978 CAAGAGCTCAAAAAGATGAATGAAATCTTCTGCAAGTCTTATGAGAAATGTAATGA 5037
 QY 953 uAlaValThrAlaThrGluThrSerGluAlaLeuArgThrGluGlu-----968
 Db 5038 ACACAAGAAAGATTCAGCATGTCGTGAAAGCTGTGAGCAAGAAACAAGAACTATAGG 5097
 QY 969 -----ValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSerAlaVa 985
 Db 5098 CAAGTTAAGAACACAGAGGCAACAGAGAGACAGAAAGCAAGTTGCAAGAGCTGA 5157
 QY 985 lSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGluValGlu 1005
 Db 5158 GCAGAAATGTGAGAAATGAAGAAAGATGAGAAAGTTT-----GCTTAATCTTAACA 5211
 QY 1005 uSerGlyValLeuAspThrGlnGluGluGluArgGlnThrGlnAlaIleLeuGlnAlaVa 1025
 Db 5212 GCAGAAATCTTAGAGCTGGAAGAAAGATGACCGGCTTAGGAGAGGTACACCTGCG 5271
 QY 1025 lAlaAspLysValLysGluGluSerGlnValProAlaThrGlnThrAlaGlnArgThrGlu 1045
 Db 5272 AGGAGATACAGTAAAGG-----GTATGGAACACATCTTCTTCCAA 5316
 QY 1045 ySerLysAlaLeuGluLysValGluGluValGluGluLysPseSerGluValAlaLeuAlaSerGlu 1065
 Db 5317 TGCCAGCATGAAGAGAAAGCTTGAAGAGGTCAAAATGAGATGTAAGAACCTTCTTAAGA 5376
 QY 1065 u-----LysGluLysAspValMetProLysGlyProValGlnGluAlaGlu 1080
 Db 5377 GTTCACTCTTAATGCTGAGAGAAAGCTCTTAAGTGA--GAGGTCAAGAT----- 5428
 QY 1080 yAlaGlnHisLeuAlaGlnGlySerGluThrGlyGlnAlaThrProGluSerLeuGluVal 1100
 Db 5429 -TTAAAGCATTCAGATGAAGAGTAATGTATCTTAACAAGTAACTAGAGGCCACGAGAGA 5487

QY 1100 lProGluValThrAlaAspValAspHisValAlaThrCysGlnVal-----llely 1117
 Db 5488 ACATGATAACCAACGAATCTACTGAGAGGGAACACAGCTATACAGGTGAGACTGA 5547
 QY 1117 sLeuGlnGlnLeuMetGluGlnAlaValAlaProGluSerSerGluThrLeuThrAspSe 1137
 Db 5548 AGAGCAAGACTCTCTGAGTATGAGCACACAGACCTACATGTTTCAAGATCGTTCATCAGC 5607
 QY 1137 rGluThrAsnGlySerThrProLeuAlaAspSerAspThrAlaAspGlyThrGlnGlnAs 1157
 Db 5608 GAAGAGT-----GCCAACCTCTGTGAAGTAAGGATTTTCAGCTCACATGATGAATTA 5661
 QY 1157 p-----GluThrIleAspSerGlnAspSerLysAlaThrAlaAlaValargGlnSe 1174
 Db 5662 TAATCTACCTACAGCAGATTTAGCTCAAGAAAGAAATGCTGGATAGAGGAGGAGAA 5721
 QY 1174 rGlnValThrGluGluGluAlaAlaThrAlaGlnLysGluGluProSerThrLeuProAs 1194
 Db 5722 GCAGAAAAACAAGGATTTAGCCAGACTTTAGAAATGAGAAA---AATACCTTACTGAG 5778
 QY 1194 nAsnValProAlaGlnGluGlnHisGlyGluGluProGlyArgAspValLeuGluProTh 1214
 Db 5779 TCAGATATCAACAAGGAT-----GGTGAA-----AlaAlaValProVa 5814
 QY 1214 rGlnGlnGluLeuThr-----AlaAlaValProVa 1225
 Db 5815 TCAGGAGGAGTAAACCAAAATGAACCTGTTAAATCAAGAAATCAAGAAAGACTCTCCAG 5874
 QY 1225 lLeuAlaLysThrGluValGlyGlnGluGlyGluValAspTrpLeuAspGlyGluLysVa 1245
 Db 5875 AGTTACCAAACTAAGGAGACAGCAGAGAGAGAAAGATGATTTGGAAGAGAGGCTTAT 5934
 QY 1245 lLysGluGlnGluValPheValHisSerGlyProAsnSerGlnLysAlaAlaAspVa 1265
 Db 5935 GAATCAATTAGCAGAACTT-----AATGGAAGCATTGGGAATTTACTGTCAGGATGT 5985
 QY 1265 lThr-----TyrAspSerGluValMetGlyValAlaGlyCy 1277
 Db 5986 TACAGATGCCAAATAAAAAATGAGCTATTGGAAATCTGGAATGAAGAACCTTAAAAAGTG 6045
 QY 1277 sGlnGluLysGluSerThrGluValGlnSerLeuSerLeuGluGluGlyGluMetGluTh 1297
 Db 6046 TGTGAGTGAATTGGAGAGAAAGACGAGCTAGTCAAGCAAAAGAACTAAGGTGGATC 6105
 QY 1297 rAspValGluLysGluLysArgGluThrLysProGluGlnValSerGluGluGlyGluGl 1317
 Db 6106 AGAAATACGAAGAGATATTGGAGAAATACAGGTGCTCAGAAAGAACCCGGAATAA 6165
 QY 1317 nGluThrAlaAlaProGluHisGluGlyThrTyrGlyLysProValLeuThrLeuAspMe 1337
 Db 6166 A-----AGCCATGCAAGAGAACTTCAGGAACCTGTTTAA 6198
 QY 1337 tProSerSerGluArgGlyLysAlaLeuGlySerLeuGlyGlySerProSerLeuProAs 1357
 Db 6199 AGAAAAACAACAGAGTAAGCAGCTA-----CAGAGGACTGCACTCAGGTA 6246
 QY 1357 pGlnAspLysAlaGlyCysIleGluValGlnValGlnSerLeuAspThrThrValThrGl 1377
 Db 6247 TCAAGAGAAAAATTAGTCTCTGGAGAGAACTGTTAAAGCTCTAGAAATTTGTTCAAACTGA 6306
 QY 1377 nThrAlaGluAlaValGluLysValIleGluThrValValIleSerGluThrGlyGluse 1397
 Db 6307 ATCTCAAAAAGATTTGGAATAACCAAGAAATA-----CTGGCTCAAGCAGTTGAACA 6360
 QY 1397 rProGluCysValGlyAlaHis-----LeuLeuProAlaGluLysSe 1411
 Db 6361 CCGCAAAAGGCACAGCAAGTAATAGCTTCAAAAGTCTCTAGATGACACTCAAG 6420
 QY 1411 rSerAlaThrGlyGlyHisThrProLeuGlnHisAlaGluAspThrValProLeuGlyPr 1431
 Db 6421 TGAAGCAGCA-----AGGTCCTAGCAGACAACTCAAGTTGAAAA 6462
 QY 1431 oGluSerGlnAla-----GluSerIleProIleValThrProAlaProGluSerTh 1449

Db 6463 GGAACCTTCAGTCAATAAAGATTCAGTTAAAGCCGATGAACAAAGGATGAA----- 6517
 QY 1449 rLeuHisProAspLeuGlnGlyGluIleSerAlaSerGlnArgGluArgSerGluGluGl 1469
 Db 6518 -----GATCTTGACGGAAGACTGGAACAGCAGCAGAGAGAGACCACTGAAGAGAA 6567
 QY 1469 uAspLysProAspAlaGlyProAspAlaAspGlyLysGluSerThrAlaIleGluLysVa 1489
 Db 6568 GAAGATATGCAAGAGAAACTGGATCTTTGGCAGAGAAAAAGTCCACTTGGAGAGAGAC 6627
 QY 1489 lLeu-----LysAlaGluProGluIleLeuGluLeuGluSe 1501
 Db 6628 AATTGGAGAGATTCAGGTACTTTGAACAAGAAAGACAAAGGAGTTCAGCAACTTCAGGA 6687
 QY 1501 rLysSerAsnLysIleVal 1507
 Db 6688 AAACCTGGACAGTACTGTG 6706
 RESULT 15
 US-09-724-676A-7240
 ; Sequence 7240, Application US/09724676A
 ; GENERAL INFORMATION:
 ; APPLICANT: Compugen LTD
 ; TITLE OF INVENTION: Variants of alternative splicing
 ; FILE REFERENCE: 129181.4 Compugen
 ; CURRENT APPLICATION NUMBER: US/09/724,676A
 ; CURRENT FILING DATE: 2000-11-28
 ; NUMBER OF SEQ ID NOS: 97222
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 7240
 ; LENGTH: 8189
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-724-676A-7240
 Alignment Scores:
 Pred. No.: 1.98e-09 Length: 8189
 Score: 415.00 Matches: 316
 Percent Similarity: 38.17% Conservative: 305
 Best Local Similarity: 19.42% Mismatches: 669
 Query Match: 5.14% Indels: 337
 DB: 5 Gaps: 60
 US-09-902-432-4 (1-1596) x US-09-724-676A-7240 (1-8189)
 QY 26 LeuSerGly-----HisGlyProAlaAlaGluAlaSerGlyAlaAla----- 39
 Db 2404 CTCCTCAGGTGAAGGAACCTTAGCATGGTGAACAGAAATTGAGGGCTCAGGTAAAGCAACTGGA 2463
 QY 40 GlyAspPro-AlaAspAlaAspProAlaThrLysLeuProGlnLysAsnGlyGlnLeuSe 59
 Db 2464 AATGAACCTTGCAAGAGCAGAAAGGCAAGAGACCTTGATTATGAAGC---CAAACTGC 2520
 QY 59 rSerValAsnGlyValAlaGluGlnGlyAspValHisValGlnGlnGluAsnGlnGluGl 79
 Db 2521 CCATGCAACCTCTCCTCACTGAACAG-----ATCCATAGTCTCAGCATAGAAGCAAAATC 2574
 QY 79 yGlnGluGluGluValValAspGluAspValGlyGlnArgGluSerGluAspValArgGl 99
 Db 2575 TAAAGATGTGAATAAT-----GAAGTTTTACAGATGAACCTGGATGATGTGCAG-- 2623
 QY 99 uLysAspArgValGluGluMetAlaAlaAsnSerThrAlaValGluAspIleThrLysAs 119
 Db 2624 -----CTTCAGTTTTCTCAGCAGAGACTACCTGTATAGAGAGCTGCAAGGCCA 2670
 QY 119 pGlyGlnGluGluThrSerGluIleIle-----GluGlnIleProAlaSerGluAs 136
 Db 2671 GCTGCAAAATTAAGAAAGTGAAGTGTCTGAGGGGGCAGAACCGTGTAAAGGCATATCTCAAG 2730
 QY 136 nAsnValGluGluMetValGlnProAlaGluSerGlnAlaAsnAspValGlyPhelyLy 156
 Db 2731 TAAAGTGAAGAACTGTCCAGGCTCTTTTCACAGAGGAAGAACTTGAATA----- 2779

QY 136 sValPheLysPheValGlyPheLysPheThrValLysLysAspLysAsnGlnLysSerAs 176
 Db 2780 -----ACAAAAATGGAT----- 2791
 QY 176 pThrValGlnLeuLeuThrValLysLysAspGlnGlnGlnValAlaGlnAlaSerValG 196
 Db 2792 -----CAGCTCTTACTAGAGAAAAAGAGATGTGGAAACCCCTCCAAACCAACCTCCA 2844
 QY 196 ValAlaGlyPheLysGlnGlnLysProSerValGlnThrAlaValGlnLysLysSerLysG 216
 Db 2845 GGAGAGAGATCAACAAGTACAGAAATCAGCTTACTAGTACAGTGAAGAAATGGTTCAGCT 2904
 QY 216 uSerGlnLeuLysGlnSerThrGlnLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 236
 Db 2905 TATGTAAGAGAAATTTCTCTGGGGGTGAAATTAAGACTCTTAAAGAA----- 2953
 QY 236 rThrGlnLysProLeuGlnAlaGlnLysSerAspGlnAlaGlnGlnGlnGlnGlnGlnGln 256
 Db 2954 -----CAGCTAAATTTTATTTCAGAGCTGAGAGCAAAAAAGACAGGTGAGAAAGA 3009
 QY 256 uGlnGlnGlnLysGlnGlnLysGlnLysProThrLysSerProGlnSerProSerProVa 276
 Db 3010 TATGTAAGTTCTCTGCGCTTAAACAAATATATGATGATGATGATGATGATGATGATGAT 3069
 QY 276 LysSerGlnLysThrSerSerPheLysLysPheThrLysGlnGlnGlnGlnGlnGlnGln 296
 Db 3070 AAGTAAAGAAAGAA-----CTTCAGCATGAAATTTGACCTTCTGTA 3108
 QY 296 GlyLysLysThrSerPheLysLysSerLys-----GlnLysAsp 308
 Db 3109 GAAAGAAATGAGCAGAGAAAGAAAGAAAGCTCCAGCAGCTTTATTAACGAAAGAGAGCT 3168
 QY 308 pLeuGlnThrAlaGlnLysArgLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 328
 Db 3169 TCTGCAAGAGTCAAGTGAATTTGAGAGAAATTAAGCCAACTTGAAGAGTAAATCTAAGA 3228
 QY 328 sGlnLysThrGlnProAlaSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 348
 Db 3229 AGAAATCCCACTCAGTGAAGCTGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3288
 QY 348 gLeuSer-----AlaAspLysGlnLysValGlnLeuProLeuGlnLysAspGln 363
 Db 3289 ATACTAGAAATATGTGTGACTTCTAAGTCCAGAGAAATGAATTTATTTAAACAGAGC 3348
 QY 363 nValGlnAspLeuGlnAla-----SerSerGlnGlnLysCysAl 376
 Db 3349 AATATCTGAGAAAGAGAGTGAATCTACAGCATATAAGAGAGATTTGAGAGAGAG----- 3403
 QY 376 aProLeuAlaThrGlnValPheAspGlnLysMetGlnAlaLysGlnGlnValAlaAlaGln 396
 Db 3404 -----CTGGCAGCTGAA-----GAGCAATTTCCAGGCTCTGTCTCAAA 3441
 QY 396 uValHisValSerThrValGlnLysThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 416
 Db 3442 GATGAAATCAAGCTTCCAAAGATTAACAAACCA----- 3475
 QY 416 uGlnGlnValValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 436
 Db 3476 -----ATAGATTGCTCCAGCAGAAATACGTGAAAA 3507
 QY 436 oGlnGlnValProGlnGlnAlaGlnProAlaGlnGlnLysLeuLysSerArgGlnMetCys 456
 Db 3508 CCAAGCAATTAATCCAGAG-----TTATCAACAAGTAACACGGA 3546
 QY 456 sValSerGlnLysAspHisThrGlnLeuThrAsp-----LeuSerProGlnGln 472
 Db 3547 TGCAGACTGATGGGACTCCGTAGCACTTTAAAGAGAAACAGTGTATTAAGTCCACCTTG 3606
 QY 472 uLysThrLeuProLysHisProGlnGlnLysLysValSerGlnValGlnMetLeuSerGln 492
 Db 3607 TACAGGTAGTAGAAACACTGGAAACCAAGAACTAAGAAAGAAATACTGCGCTTGAATA 3666

QY 492 nGlnArgLysValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 512
 Db 3667 AGAAAGAGAGCACTTCAAAAG 3726
 QY 512 sLysLeuSerGlnLysLysGlnLysGlnLysArgGlnGlnGlnGlnGlnGlnGlnGlnGln 532
 Db 3727 TAAAG 3771
 QY 532 yGlnLysGlnHisLysLysThrGlnLysProGlnLysSerAlaAspGlnGlnGlnGlnGln 552
 Db 3772 GCACAG 3831
 QY 552 r-----SerAl 554
 Db 3832 TGAGAAATTTGAG 3891
 QY 554 aSerSerProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 574
 Db 3892 AAAAGTCCAG 3951
 QY 574 oGlnAspGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 584
 Db 3952 TTTATCAAG 4011
 QY 585 -----SerAspGlnGlnLysLysArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 599
 Db 4012 AGACTGCGCTTCTCTCTGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4071
 QY 599 rPheLysLysMetValThrProLysLysArgValArgArgProSerGlnSerLysAspGln 619
 Db 4072 GATTAAAGGCCAGCTG-----AAGGA 4092
 QY 619 uGlnGlnLysGlnLysValLys-----SerAlaThrLeuSerSerThrAspSerThrVal 638
 Db 4093 AATAG 4152
 QY 638 rGlnMetGlnAspGlnLysLysThrValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 658
 Db 4153 TAAAGAAATCAG 4212
 QY 658 gArgValAspThrSerValSerThrGlnAlaLeuLysCysValGlnSerSerLysLysArg 678
 Db 4213 CGAGAGCTTAAG 4272
 QY 678 gAlaArgLysAlaSerSerSerAspAspGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 698
 Db 4273 ATTTG-----GAAAGCAGCCAACTAACAATTTGCGCTGAG 4308
 QY 698 rHisArgAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 718
 Db 4309 ACATCTTAAG 4359
 QY 718 rThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 735
 Db 4360 CAAAG 4419
 QY 735 ySerProSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 755
 Db 4420 CACTAAATATCAAG 4477
 QY 755 oArgLysLysSerLysSerLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 775
 Db 4478 -----CAGCTAAGAAATCAAGCCAAAGAGATGTGAGAGAGATTAAGACAGT 4524
 QY 775 uSerThrGlnLys-----GlnProSerArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 788
 Db 4525 AAGAGTGAAGCTTTGTAATTAAGCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4579
 QY 788 LserLysLysPheLysLysProGlnArgArgLysLysArgLysLysLysLysLysLysLys 808
 Db 4580 -----AGAGCAAG 4614
 QY 808 nAlaThrValGlnAspSerGlnProValGlnLysAsnGlnLysAspProAsnValProAl 828

Qy 1489 lleu-----LysAlaGluProGluIleLeuGluLeuGluSe 1501
 Db 6628 AATTGGAGAGATTCAAGTTACTTGAACAAAGAAAGCAAGTTCAAGCAACTTCAGGA 6687
 Qy 1501 rlySerAsnLysIleVal 1507
 Db 6688 AAACCTTGGACAGTACTGTG 6706

Search completed: December 13, 2002, 05:34:53
 Job time : 2025 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: December 12, 2002, 14:13:30 ; Search time 14973 Seconds
(without alignments)
11973.104 Million cell updates/sec
Title: US-09-902-432-3
Perfect score: 6160
Sequence: 1 cggcgccggcgagtagaag.....aaataaacattattgactac 6160

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0
Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_cm.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vt.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	6160	100.0	6160	10	RNU23146	U23146 Rattus norv
2	5204.2	84.5	5236	10	RNU41453	U41453 Rattus norv
3	4831.4	78.4	5134	6	AR071282	AR071282 Sequence
4	4559.6	74.0	6197	10	AF326228	AF326228 Mus muscu
5	4558	74.0	6195	10	AB020886	AB020886 Mus muscu
6	4441.8	72.1	178954	2	AC115647	AC115647 Rattus no
7	4304	69.9	5902	10	AF326230	AF326230 Mus muscu
8	2179.6	35.4	6614	6	AX405857	AX405857 Sequence
9	2144.4	34.8	6608	6	AX329766	AX329766 Sequence
10	2144.4	34.8	6608	6	AX410793	AX410793 Sequence
11	2144.4	34.8	6608	9	HSU81607	U81607 Homo sapien
12	2136.8	34.7	6297	9	AF001504	AF001504 Homo sapi
13	2118.8	34.4	6605	6	AR002603	AR002603 Sequence
14	2089.6	33.9	6287	9	AB003476	AB003476 Homo sapi
15	1906.4	30.9	104939	9	AL590413	AL590413 Human DNA
16	1906.4	30.9	171158	2	AC023201	AC023201 Homo sapi
17	923.8	15.0	102822	2	AC097731	AC097731 Rattus no
18	911.8	14.8	178954	2	AC115647	AC115647 Rattus no
19	681.2	11.1	129179	2	AC117205	AC117205 Mus muscu
20	651	10.6	102822	2	AC097731	AC097731 Rattus no
21	550	8.9	563	6	AX401547	AX401547 Sequence
22	466.4	7.6	1027	9	BC000188	BC000188 Homo sapi
23	466.4	7.6	1027	9	BC022814	BC022814 Homo sapi
24	416	6.8	8439	5	AF308810	AF308810 Xenopus l
25	313.2	5.1	1981	9	HUMGRAVIN	M96322 Human gravi
26	232.8	3.8	1058	11	G06673	G06673 human STS W
27	163.4	2.7	129179	2	AC117205	AC117205 Mus muscu
28	162.2	2.6	178	6	AX401799	AX401799 Sequence
29	162.2	2.6	178	10	RNU75404	U75404 Rattus norv
30	160.6	2.6	696	9	HUMZD39H03	AF086250 Homo sapi
31	131.8	2.1	386	11	G21790	G21790 human STS W
32	108.6	1.8	179145	9	AC009785	AC009785 Homo sapi
33	107.8	1.7	195690	10	AC098726	AC098726 Rattus no
34	104.8	1.7	177881	2	AC116238	AC116238 Mus muscu
35	104.6	1.7	174990	10	AC027653	AC027653 Mus muscu
36	101.6	1.6	35848	2	AC011553	AC011553 Homo sapi
37	101.2	1.6	33082	2	AC091365	AC091365 Rattus no
38	101	1.6	193363	2	AL627104	AL627104 Mus muscu
39	99.8	1.6	7218	6	I66494	I66494 Sequence 14
40	99.8	1.6	68100	2	AC105787	AC105787 Rattus no
41	99.6	1.6	196194	2	AC104920	AC104920 Mus muscu
42	99.2	1.6	206497	10	AC084389	AC084389 Mus muscu
43	98.8	1.6	45632	2	AC098158	AC098158 Rattus no
44	98.6	1.6	125816	10	AL591127	AL591127 Mouse DNA
45	98.2	1.6	194334	2	AC101527	AC101527 Mus muscu

ALIGNMENTS

RESULT 1
RNU23146
LOCUS Rattus norvegicus mitogenic regulation ssecks (322) gene, complete cds.
DEFINITION RNU23146 6160 bp DNA linear ROD 19-OCT-1995
ACCESSION U23146
VERSION U23146.1 GI:1022682
KEYWORDS Rattus norvegicus.
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
REFERENCE 1 (bases 1 to 6160)
AUTHORS Lin,X., Nelson,P.J., Frankfort,B., Tomblor,E., Johnson,R. and

TITLE Gelman, I.H.
Isolation and characterization of a novel mitogenic regulatory gene, 322, which is transcriptionally suppressed in cells transformed by src and ras

JOURNAL Mol. Cell. Biol. 15 (5), 2754-2762 (1995)

MEDLINE 95257957

PUBMED 7739556

REFERENCE 2 (bases 1 to 6160)

AUTHORS Gelman, I.H.

TITLE Direct Submission

JOURNAL Submitted (21-MAR-1995) Irwin H. Gelman, Microbiology, Mount Sinai School of Medicine, One Gustave L. Levy Place, New York, NY 10029, USA

COMMENT On Oct 19, 1995 this sequence version replaced gi:733449.

FEATURES

source location/Qualifiers

1..6160

/organism="Rattus norvegicus"

/strain="fisher"

/db_xref="taxon:10116"

/clone="13.2.2"

/cell_line="Rat-6"

/cell_type="fibroblast"

/tissue_type="embryo"

/clone_lib="RI-a (Robert Kraus, Mount Sinai School of Medicine)"

1..32

33..5044

/gene="322"

33..4823

/gene="322"

/function="possible mitogenic regulatory gene"

/note="Src Suppressed C Kinase Substrate."

/citation=[1]

/codon_start=1

/evidence=experimental

/product="SSECKS"

/protein_id="AA79517.1"

/db_xref="GI:1022683"

/translation="MGAGSSTGGRSPPEOPAGSDTPSELVLSGHPAABASGACDPADADPATKLPQKNGOLSVNGVAGQDVHVGQENQGOEEVVDVGORSESDVEKORVEEMANSTAVEDITDKQDEETSEIIEQIPASNNVGVQVPAESQANDVPFKVFPFVGFPTVKDKNEKSDTVQLTVKKDEBGEAEVAGDHPQEVETAVGSAKESLKGSTBEKQGLKQKQSGSTPILOAESQAAEBAKQDEKQKQKPEKSPSPVNSSTTSFKKFTFGWAGMKRTKTSFKSKEDDLTAERKEQAEKDEBEKTEBAVSEQEPADDTQALISADYEKVELPLEQVGLDLSASEKCAPLATEVPEKMAHDEVVAEVHVSIVTEKTEEBQGGGAEAGVVEGTGSLPPEKLAQEVOEAPBEIMKSRMCVSGGHTQLTDLSPBEKTLPHBEGIVSEVEMLSQERIKVQSPKLFPSGLKIKSGKQKQKGGGDEBGEYOHITHESPESADDEQKSGSSASSPEPEETCLKGPLPAPQDEABEGTTSDEKKEKREGITTPWASPKKWTTPKKVRPSPSEKRELEKVKSAATLSTDSIVSEMDEVKTVGEBOQPEEPKRRVDTSVSEWALICVGSKKARAKSSSDDEGPRTLGGDSHRAEASKQKEKADSVASTQEOQAGSSSPBAGSPBSGEGVSTWESFKRLVTPRKSKSKLEKAKEDSVLEQDSTEIPEPSRESEVWSIKFPGRRKKRADGQEOAEVDSGVEINEDDPNVPVAVPSEVAVAREMEAGNTELPOLLGAVVSEBELSKTLVHTVSVAIVDGTAVTSVERSPMSIATVTEPLHETAGEMPVEEVTAKDITAEETPVLTQTLPEKQAHDMNTSEVDTPSEATATFETSEALTEVETAEESQAEETDMVSAVQLTQSPDTEEATPVQVEGVLDTSEEEKQTAIIQAVDVKVEESQVATQTVORTGSKALEKVEVEEDSEVLAEKEDVMPKQPVQAEAGHAAQKSEGOATPESLPEPVTADVAVATCOVAKILOLMEOAAVAPSSSTLIDSETNGSTPLADSDTADTQOODETTDSQSKATAVROSOVTEBEATATQKPESTLPPNNVPAOEHGEPRGDVLEPTQOELTAAPVPIATKTEVQGEVVMLOKEKKEQOEVAHVHSGPMSOKAADVIVDEVMVAGQOEKESTEVOSLSLEGEMETDVKEKRETKPEVSEEGSOETAAPHEHGVGKPVLTLDMPSSERKALGSLGSPSLPDQDKGCTLEVOQSLDTVTOTAAEAEKVEIVETVSETGSPCEVAGAHLLPAKKSATGQHWTLQHAEDVTPAGPSQASIPITVPAPESTLHPDQGEISASORRSEEDKPDGADKESVIAIKVKAEPETLFLSKSNKIVINVTOTAVDOPARTETTABETHAYDSOQVPCRLDSRPNKCMWKMKDKAKMHPVQPPREDLQVLTVLEMAQPKCRLPRLQKAVSK"

misc_signal 1581..1590

/gene="322"

/note="encodes nuclear localization signal; NLS"

/citation=[1]

1797..1811

/gene="322"

/note="encodes nuclear localization signal; NLS"

misc_signal

misc_signal /citation=[1]

2061..2073

/gene="322"

/note="encodes nuclear localization signal; NLS"

misc_signal /citation=[1]

2301..2313

/gene="322"

/note="encodes nuclear localization signal; NLS"

misc_binding /citation=[1]

4230..4358

/gene="322"

/note="encodes zinc-finger domain"

3' UTR /bound_moiety="DNA, RNA or protein binding domain"

misc_signal 4816..6160

5032..5044

/gene="322"

/note="encodes nuclear localization signal; NLS"

polyA_signal 5760..5765

polyA_signal 6136..6141

polyA_site 6160

BASE COUNT 1853 a 1409 c 1753 g 1145 t

ORIGIN

Query Match 100.0%; Score 6160; DB 10; Length 6160;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 6160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGCGCGCGGAGTATAGAGCCACTGACCATGAGCGGAGGAGGAGTTCACCGAGAGC 60

DB 1 CGGCGCGCGGAGTATAGAGCCACTGACCATGAGCGGAGGAGGAGTTCACCGAGAGC 60

QY 61 GGAGCCCCGAGCAGCCGCGGGGAGCAGCAGCCGAGCAGCTGGTCTCAGTGGCCATG 120

DB 61 GGAGCCCCGAGCAGCCGCGGGGAGCAGCAGCCGAGCAGCTGGTCTCAGTGGCCATG 120

QY 121 GGGCCGAGCTGAAGCCTCGGGAGCACTGAGACCCCCCGGACGCGGACCCCGCCACCA 180

DB 121 GGGCCGAGCTGAAGCCTCGGGAGCACTGAGACCCCCCGGACGCGGACCCCGCCACCA 180

QY 121 GGGCCGAGCTGAAGCCTCGGGAGCACTGAGACCCCCCGGACGCGGACCCCGCCACCA 180

DB 121 GGGCCGAGCTGAAGCCTCGGGAGCACTGAGACCCCCCGGACGCGGACCCCGCCACCA 180

QY 181 AGCTCCACAGAAATGCGCAGCTGTCTTCTGCAACGCGGTAGTGAACAAGAGATG 240

DB 181 AGCTCCACAGAAATGCGCAGCTGTCTTCTGCAACGCGGTAGTGAACAAGAGATG 240

QY 241 TCCATGTCGAAAGAGAAACCAAGAGGGGCGAGAGAAAGTGTGATGAGATGTTG 300

DB 241 TCCATGTCGAAAGAGAAACCAAGAGGGGCGAGAGAAAGTGTGATGAGATGTTG 300

QY 301 GACGCGGAGATCGAAGATGTGAGAGAAAAGCCGAGTTGAAGAAATGGCGGCAACT 360

DB 301 GACGCGGAGATCGAAGATGTGAGAGAAAAGCCGAGTTGAAGAAATGGCGGCAACT 360

QY 301 GACGCGGAGATCGAAGATGTGAGAGAAAAGCCGAGTTGAAGAAATGGCGGCAACT 360

DB 301 GACGCGGAGATCGAAGATGTGAGAGAAAAGCCGAGTTGAAGAAATGGCGGCAACT 360

QY 361 CCAGAGCTGTGAAGATATCACAAGAGTGGCGAGGAGACATCGAATAATTTGAC 420

DB 361 CCAGAGCTGTGAAGATATCACAAGAGTGGCGAGGAGACATCGAATAATTTGAC 420

QY 421 AGATCCCTGCTTCAGAAAACAATGTGAGAAATGTACAGCTGTAGTCCAGGCTA 480

DB 421 AGATCCCTGCTTCAGAAAACAATGTGAGAAATGTACAGCTGTAGTCCAGGCTA 480

QY 481 ATGATGTTGGCTTCAAGAAAGTATTTAAATTTGTTGTTTAAATTCACGGTGAAGAG 540

DB 481 ATGATGTTGGCTTCAAGAAAGTATTTAAATTTGTTGTTTAAATTCACGGTGAAGAG 540

QY 541 ATAAATATGAATACTGATCTGTCCACATCTGTCACATGAAGATTAAGCGCAAG 600

DB 541 ATAAATATGAATACTGATCTGTCCACATCTGTCACATGAAGATTAAGCGCAAG 600

QY 601 GGGCAGAGCCTCTGTGCGAGCTGAGACCAACGAGGCCAGTGTGAGAGCTCGCTCG 660

DB 601 GGGCAGAGCCTCTGTGCGAGCTGAGACCAACGAGGCCAGTGTGAGAGCTCGCTCG 660

Db	2821	CCCAGACGTTACCAAGAGGGTTAAAGATGTCCTCATACGACATGGTCACTCACTGAAGTGAATT	2880
Qy	2881	TCACCTCAGAAAGCTGTGACAGCCACAGAGACCTCAGAGGCTCTCCGTACTGAAGAATT	2940
Db	2881	TCACCTCAGAAAGCTGTGACAGCCACAGAGACCTCAGAGGCTCTCCGTACTGAAGAATT	2940
Qy	2941	CCGAAGCATCGGGGGGCCGAAGAGACCAACATGTTGCCGATTTTCCAGCTGACTG	3000
Db	2941	CCGAAGCATCGGGGGGCCGAAGAGACCAACATGTTGCCGATTTTCCAGCTGACTG	3000
Qy	3001	ACTCCCAAGACACCAAGAGAAAGCCACCACCTCAGAGGAGTGAAGTGTGTGCTAG	3060
Db	3001	ACTCCCAAGACACCAAGAGAAAGCCACCACCTTACAGAGGTAGAGAGTGTGTGCTAG	3060
Qy	3061	ATACAGAAAGAAAGAGAGCGCCAGACCGCAGGCGCTCTCCAAAGCCGTTGCAGCAAGAGTGA	3120
Db	3061	ATACAGAAAGAAAGAGAGCGCCAGACCGCAGGCGCTCTCCAAAGCCGTTGCAGCAAGAGTGA	3120
Qy	3121	AAGAGAGTCCAGGTGCTTGCAACCTGACACTGTGCAGAGAAACGGGGTCAAAAGCATCTGG	3180
Db	3121	AAGAGAGTCCCAAGTGTCTTGCAACCAGACTGTGCAGAGAAACGGGGTCAAAAGCATCTGG	3180
Qy	3181	AGAAAGTTGAGAGAGGTAGAGAGAGACTCCGAATGTGGCTTCGAGAAAGAGAGAGCG	3240
Db	3181	AGAAAGTTGAGAGAGGTAGAGAGAGACTCCGAATGTGGCTTCGAGAAAGAGAGAGCG	3240
Qy	3241	TTATGCCGAAAGGACCCGCTGCAGAGAGCTTGAGCTTGTCACAGGCGCTCTGAGA	3300
Db	3241	TTATGCCGAAAGGACCCGCTGCAGAGAGCTTGAGCTTGTCACAGGCGCTCTGAGA	3300
Qy	3301	CTGACACAGGCTACTCCAGAGAGCCTTGAAATCTTGAAAGTCAAGGCACATGTGACACTG	3360
Db	3301	CTGACACAGGCTACTCCAGAGAGCCTTGAAATGTTCTTAAAGTCAAGGCATGTGACACTG	3360
Qy	3361	TGCGCACGTGCCAGGTTATACAGCTCCAGAGGTGATGGAACAGGCGGTGGCCCTGAGT	3420
Db	3361	TGCGCACGTGCCAGGTTATACAGCTCCAGAGGTGATGGAACAGGCGGTGGCCCTGAGT	3420
Qy	3421	CATCCGAAACCTTACACAGACAGTGAACAATGGAAGCACTCCCTTACAGATTCAGACA	3480
Db	3421	CATCCGAAACCTTACACAGACAGTGAACAATGGAAGCACTCCCTTACAGATTCAGACA	3480
Qy	3481	CTGCAGATGGGACACAGCAGAGATGAACCATGACACCGACAGACAGTAAAGCCACTGACG	3540
Db	3481	CTGCAGATGGGACACACAGAGATGAACCATTTGACACCGACAGACAGTAAAGCCACTGACG	3540
Qy	3541	CTGTACAGGCACTGCACAGGTACACGAABAAGGCGGTACTGCTCAGAAAGAGAGCCTT	3600
Db	3541	CTGTACAGGCACTGCACAGGTACACGAABAAGGCGGTACTGCTCAGAAAGAGAGCCTT	3600
Qy	3601	CGACACTACCTTAATATGTTCCAGCCCGAGAAACAATGGGGAGAAACAGGAAAGATG	3660
Db	3601	CGACACTACCTTAATATGTTCCAGCCCGAGAAACAATGGGGAGAAACAAGGAAAGATG	3660
Qy	3661	TTCTTGAACCTTACACAGCAGAGCTTACTGCTCAGCCCGTCTTGCGCAAAAGACTG	3720
Db	3661	TTCTTGAACCTTACACAGCAGAGCTTACTGCTCAGCCCGTCTTGCGCAAAAGACTG	3720
Qy	3721	AGGTGGGTCAAGAGGTGAGTTGACTGTGTGATGGAAGAAAGTCAAGAAAGAACAGG	3780
Db	3721	AGGTGGGTCAAGAGGTGAGTTGACTGTGTGATGGAAGAAAGTCAAGAAAGAACAGG	3780
Qy	3781	AGGTTTGTGACACTCTGAGACCCAAACGTCAAAAGGCTGTATGTGACATTTGACAGTG	3840
Db	3781	AGGTTTGTGACACTCTGAGACCCAAACGTCAAAAGGCTGTATGTGACATTTGACAGTG	3840
Qy	3841	AAGTGATGGAGTGGCCGGGTGTCCAGAGAAAGAGAGTACTGAAGTGCAGAGCTTTAGCC	3900
Db	3841	AAGTGATGGAGTGGCCGGGTGTCCAGAGAAAGAGAGTACTGAAGTGCAGAGCTTTAGCC	3900
Qy	3901	TGGAGAGGAGAGATGGAAACTGACGTTGAAAGAGAAAGGAGACAAAGCCACAGAGC	3960
Db	3901	TGGAGAGGAGAGATGGAAACTGACGTTGAAAGAGAAAGGAGACAAAGCCACAGAGC	3960

QY	3951	AAAGACTGAAGAAAGGTGAGCAGAGAAACACGCCGCTCTGACATGAAGAAACCTAACGCA	4020
Db	3961	AAAGTAGAGAAAGAAAGGTGAGCAGAGAAACGCCGCTCTGACATGAAGAAACCTAACGCA	4020
QY	4021	AGCCAGTCTTGACACTTGACATGCCAGCTCAGAGAGGGGGGAGCACTGGGAAAGCCTTG	4080
Db	4021	AGCCAGTCTTGACACTTGACATGCCAGCTCAGAGAGGGGGGAGCACTGGGAAAGCCTTG	4080
QY	4081	GAGGAAGCCCTTCTCTCCAGACCAAGACAAAGCAGGTTGCATAGAGTTCAAGTTCAA	4140
Db	4081	GAGGAAGCCCTTCTCTCTCCAGACCAAGACAAAGCAGGTTGCATAGAGTTCAAGTTCAA	4140
QY	4141	GCCCTGACACAAACAGTCACTCAACACGACGAAGCTGTGAAAAAGGCTCAAGAAACGGTTG	4200
Db	4141	GCCCTGACACAAACAGTCACTCAACACGACGAAGCTGTGAAAAAGGCTCAAGAAACGGTTG	4200
QY	4201	TGATTTGAGAGACAGGTGAAGATGCGAGTGTGTAGTGCACTTATTAACCTGACTGAGA	4260
Db	4201	TGATTTGAGAGACAGGTGAAGATGCGAGTGTGTAGTGCACTTATTAACCTGACTGAGA	4260
QY	4261	AGTCTCTGCAACGGGTGGCAATTGGAACCTTCACATGACAGAGACACGGTTACCCCTGG	4320
Db	4261	AGTCTCTGCAACGGGTGGCAATTGGAACCTTCACATGACAGAGACACGGTTACCCCTGG	4320
QY	4321	GGCCTGAGTCTCAGGCGAGATTCACATCATATGTAAGTTCTTCTCTGAAAAAGCACCC	4380
Db	4321	GGCCTGAGTCTCAGGCGAGATTCACATCATATGTAAGTTCTTCTCTGAAAAAGCACCC	4380
QY	4381	TACATCTGACTCAACGAGAAATTAAGCCATCCCAAGAGAGCGATCAAGAGAAAGG	4440
Db	4381	TACATCTGACTCAACGAGAAATTAAGCCATCCCAAGAGAGCGATCAAGAGAAAGG	4440
QY	4441	ACAAAGCCAGATGCGTGTCCATAGTCTGACGGGACGAGAGATGCACATTCGAAAAAGTCC	4500
Db	4441	ACAAAGCCAGATGCGTGTCCATAGTCTGACGGGACGAGAGATGCACATTCGAAAAAGTCC	4500
QY	4501	TCAAAGCTGAACCTGAGATCTTGSAACCTTGAAAGTAAAGCAACAAATTTGTCTGAACG	4560
Db	4501	TCAAAGCTGAACCTGAGATCTTGSAACCTTGAAAGTAAAGCAACAAATTTGTCTGAACG	4560
QY	4561	TCATTCAGACAGCCGCTTGACCAATGTGGACGTACGAAACAGCCCCGAAACTCATGCTT	4620
Db	4561	TCATTCAGACAGCCGCTTGACCAATGTGGACGTACGTACGAAACAGCCCCGAAACTCATGCTT	4620
QY	4621	ATGATTCACAGACCCAGATTCCTGCATGTGACAGGCTTGACAGCAGGAGGCCAACAGATGCT	4680
Db	4621	ATGATTCACAGACCCAGATTCCTGCATGTGACAGGCTTGACAGCAGGAGGCCAACAGATGCT	4680
QY	4681	GGAACAAAAATGAAGAATGCGCAATGAAACACCCAGTGGCGGACGCCAGAGAGAGACTTGC	4740
Db	4681	GGAACAAAAATGAAGAATGCGCAATGAAACACCCAGTGGCGGACGCCAGAGAGAGACTTGC	4740
QY	4741	AAAGTCTGACCGGTTCTGGAAGGCAATGGGCTCAGGCTGGAATAATGCTTGGCGGCTTGAGT	4800
Db	4741	AAAGTCTGACCGGTTCTGGAAGGCAATGGGCTCAGGCTGGAATAATGCTTGGCGGCTTGAGT	4800
QY	4801	TGAAAAGCGCGGTGTCAAAGTAAGCAATTGAAGAGTGCTCTCTCAACCCAAAGATCCAA	4860
Db	4801	TGAAAAGCGCGGTGTCAAAGTAAGCAATTGAAGAGTGCTCTCTCAACCCAAAGATCCAA	4860
QY	4861	AGGAGCAATGCTGATGAGCGGCTCAGTCCAAAGCTTAACCCAGGCAAGAGCGCAGTGGCT	4920
Db	4861	AGGAGCAATGCTGATGAGCGGCTCAGTCCAAAGCTTAACCCAGGCAAGAGCGCAGTGGCT	4920
QY	4921	CTGGAACCTTAACCAAGAAATCCCGACACCAACGGACCAAAAGCTAACCGAGAGGGCG	4980
Db	4921	CTGGAACCTTAACCAAGAAATCCCGACACCAACGGACCAAAAGCTAACCGAGAGGGCG	4980
QY	4981	ATCCCCCAAAAAGTTGAGGTCAGAGAAAGAAATGTTCTACAGTCAAGTCAAGAGAAACA	5040
Db	4981	ATCCCCCAAAAAGTTGAGGTCAGAGAAAGAAATGTTCTACAGTCAAGTCAAGAGAAACA	5040

Qy	5041	AGGCCAGG	CAGAGAGG	AGGACCTG	CAGAGGCC	AAAGGGAG	AGCTG	GCAGAA	TCTT	AGAGATG	5100
Db	5041	AGGCCAGG	CAGAGAGG	AGGACCTG	CAGAGGCC	AAAGGGAG	AGCTG	GCAGAA	TCTT	AGAGATG	5100
Qy	5101	TTAGTTCT	CATTGT	ATCATCTG	TTAAGC	CAGAAATG	TGAACA	CAAGTGC	ACAGAA	CAAGATG	5160
Db	5101	TTAGTTGCT	CATTGT	ATCATCTG	TTAAGC	CAGAAATG	TGAACA	CAAGTGC	ACAGAA	CAAGATG	5160
Qy	5161	CTGCTGTT	TGGGAC	CTTTGAG	ACCAAG	ATTTT	CAGAGCCC	ATGAGAT	CCAGAGAC	AGGCGCG	5220
Db	5161	CTGCTGTT	TGGGAC	CTTTGAG	ACCAAG	ATTTT	CAGAGCCC	ATGAGAT	CCAGAGAC	AGGCGCG	5220
Qy	5221	TCCAAATG	ATTTT	CCACCG	ATAGAG	CACCCG	ACAATTT	CTGAGG	CTTCAT	TCGGGAGCT	5280
Db	5221	TCCAAATG	ATTTT	CCACCG	ATAGAG	CACCCG	ACAATTT	CTGAGG	CTTCAT	TCGGGAGCT	5280
Qy	5281	GCAGCT	TAACAT	TTCTCG	TTTTC	CAAGAC	TGCCCT	TTGATTT	TGCCCT	TGATGCGGCTCGG	5340
Db	5281	GCAGCT	TAACAT	TTCTCG	TTTTC	CAAGAC	TGCCCT	TTGATTT	TGCCCT	TGATGCGGCTCGG	5340
Qy	5341	TATTTCT	TAA	CAATTT	TCCTCG	TTTTC	CAAGAC	TGCCCT	TTTGA	TTTGGCCCTT	5400
Db	5341	TATTTCT	TAA	CAATTT	TCCTCG	TTTTC	CAAGAC	TGCCCT	TTTGA	TTTGGCCCTT	5400
Qy	5401	TATTT	TCGGA	TTT	AAAGTC	CTCGG	TTCT	CAAC	CTG	GAA	5460
Db	5401	TATTT	TCGGA	TTT	AAAGTC	CTCGG	TTCT	CAAC	CTG	GAA	5460
Qy	5461	ACTTCT	C	MAACTG	GAGATC	CTCCTT	ATATG	TATTT	TATG	TATGTT	5520
Db	5461	ACTTCT	C	MAACTG	GAGATC	CTCCTT	ATATG	TATTT	TATG	TATGTT	5520
Qy	5521	TCCTG	TACCT	TAT	TG	TATATTTT	TTTCT	TAAG	CTG	TAA	5580
Db	5521	TCCTG	TACCT	TAT	TG	TATATTTT	TTTCT	TAAG	CTG	TAA	5580
Qy	5581	TATATA	ACGGG	TG	CAGC	CA	TAG	CGAC	CTT	TG	5640
Db	5581	TATATA	ACGGG	TG	CAGC	CA	TAG	CGAC	CTT	TG	5640
Qy	5641	TG	CAG	CAAA	CAG	ATAC	ATTT	CTG	GC	CAAG	5700
Db	5641	TG	CAG	CAAA	CAG	ATAC	ATTT	CTG	GC	CAAG	5700
Qy	5701	GC	TT	AG	AT	CTG	TG	GGG	CTT	CT	5760
Db	5701	GC	TT	AG	AT	CTG	TG	GGG	CTT	CT	5760
Qy	5761	CAG	AA	T	AAAA	CCCC	CA	TTT	G	A	5820
Db	5761	CAG	AA	T	AAAA	CCCC	CA	TTT	G	A	5820
Qy	5821	CCCC	CTT	TT	TG	CTA	AT	TG	CA	AA	5880
Db	5821	CCCC	CTT	TT	TG	CTA	AT	TG	CA	AA	5880
Qy	5881	CT	GT	CT	CT	TA	AT	TG	CA	AA	5940
Db	5881	CT	GT	CT	CT	TA	AT	TG	CA	AA	5940
Qy	5941	AGA	AG	GA	AT	CT	GT	T	G	A	6000
Db	5941	AGA	AG	GA	AT	CT	GT	T	G	A	6000
Qy	6001	GC	AG	AT	CG	AG	CT	CT	TT	T	6060
Db	6001	GC	AG	AT	CG	AG	CT	CT	TT	T	6060
Qy	6061	TG	T	GT	T	A	T	AT	CT	G	6120
Db	6061	TG	T	GT	T	A	T	AT	CT	G	6120
Qy	6121	TG	T	G	AT	T	G	A	A	A	6160

[illegible]

Query Match 84.5%; Score 5204.2; DB 10; Length 5236;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 5223; Conservative 0; Mismatches 13; Indels 1; Gaps 1;

QY 1 CGGGCCGGCGGAGTAGAAGCACTGAGCCATGCGCGCAGGCAATTCCACCGACGAC 60
|
|
|
Db 1 CGGGCCGGCGGAGTAGAAGCACTGAGCCATGCGCGCAGGCAATTCCACCGACGAC 60
|
|
|
QY 61 GGAACCCCGAGACCCCGCGGAGCCGACACGCGGAGACCTGTGCTCAGTGGCCATG 120
|
|
|
Db 61 GGAACCCCGAGACCCCGCGGAGCCGACACGCGGAGACCTGTGCTCAGTGGCCATG 120
|
|
|
QY 121 GGCCCGCAGCTGAAGACCTCGGAGACGTGAGACCCCGCAGCGCGACCCCGCACCA 180
|
|
|
Db 121 GGCCCGCAGCTGAAGACCTCGGAGACGTGAGACCCCGCAGCGCGACCCCGCACCA 180
|
|
|
QY 181 AGCTCCCAAGAAATGCGCAGCTGTCTTCTGTCACCGGTGACTGAACAAGAGATG 240
|
|
|
Db 181 AGCTCCCAAGAAATGCGCAGCTGTCTTCTGTCACCGGTGACTGAACAAGAGATG 240
|
|
|
QY 241 TTCATGTCGAAGAGAAAACCAAGAGGGGCAAGAGAAAGTGGTATGAGATGTTG 300
|
|
|
Db 241 TTCATGTCGAAGAGAAAACCAAGAGGGGCAAGAGAAAGTGGTATGAGATGTTG 300
|
|
|
QY 301 GACGCGGAGAGTCGAAGATGAGAGAAAAGACCGAGTTGAAGAAATGGCGCCAACT 360
|
|
|
Db 301 GACGCGGAGAGTCGAAGATGAGAGAAAAGACCGAGTTGAAGAAATGGCGCCAACT 360
|
|
|
QY 361 CCACAGCTGTTGAAGATATCAAAAGATGGCAGAGAGACATCAAGAAATATTTGAAC 420
|
|
|
Db 361 CCACAGCTGTTGAAGATATCAAAAGATGGCAGAGAGACATCAAGAAATATTTGAAC 420
|
|
|
QY 421 AGATCCCTGCTTCAGAAAACAATGTGGAAGAAATGTTACAGCTTCTAGTCCGAGGCTA 480
|
|
|
Db 421 AGATCCCTGCTTCAGAAAACAATGTGGAAGAAATGTTACAGCTTCTAGTCCGAGGCTA 480
|
|
|
QY 481 ATGATGTTGGCTTCAAGAAAGTATTTAAATTTGTTTAAATTCACGCTGAAGAG 540
|
|
|
Db 481 ATGATGTTGGCTTCAAGAAAGTATTTAAATTTGTTTAAATTCACGCTGAAGAG 540
|
|
|
QY 541 ATAAATATGAAAAAGTCAGATCTGTCCAACTACTCACTGCAAGAGATGAAGCGGAG 600
|
|
|
Db 541 ATAAATATGAAAAAGTCAGATCTGTCCAACTACTCACTGCAAGAGATGAAGCGGAG 600
|
|
|
QY 601 GGGCAGAAAGCTTCTGTCGAGCTGAGAGACACACAGAGCCCGGTGTGAGACTGCGCTG 660
|
|
|
Db 601 GGGCAGAAAGCTTCTGTCGAGCTGAGAGACACACAGAGCCCGGTGTGAGACTGCGCTG 660
|
|
|
QY 661 GAGATCGACATCCAAAGAAAGTGAAGCTGAAGCAATCCACAGAGAAAGCAAGAGCACCC 720
|
|
|
Db 661 GAGATCGACATCCAAAGAAAGTGAAGCTGAAGCAATCCACAGAGAAAGCAAGAGCACCC 720
|
|
|
QY 721 TGAAGCAAGAACAGACAGACAGAAATCCCCCTTCAAGCCGAATCTGATCAAGCGGCTG 780
|
|
|
Db 721 TGAAGCAAGAACAGACAGACAGAAATCCCCCTTCAAGCCGAATCTGATCAAGCGGCTG 780
|
|
|
QY 781 AGGAAGAGCCCAAAGATGAAGGAAGAAACCAAGAGAAAGAGCCCAAGTCCCGCAG 840
|
|
|
Db 781 AGGAAGAGCCCAAAGATGAAGGAAGAAACCAAGAGAAAGAGCCCAAGTCCCGCAG 840
|
|
|
QY 841 AATCCCGCAGAGCCCGCAGCAACAGTGAAGCAACATCTTCTTCAAGAGTTCTTCACTC 900
|
|
|
Db 841 AATCCCGCAGAGCCCGCAGCAACAGTGAAGCAACATCTTCTTCAAGAGTTCTTCACTC 900
|
|
|
QY 901 ACGGTTGGCCCGCTGGCCCAAGAAAGCAGTTCAAGAAATCAAAAAGAGATGATCTG 960
|
|
|
Db 901 ACGGTTGGCCCGCTGGCCCAAGAAAGCAGTTCAAGAAATCAAAAAGAGATGATCTG 960
|
|
|
QY 961 AAATCTCCAGAAAGAGAAAGAGCAAGAGCGAGCAAAAAGTAAAGTCGAGAGAAAGAA 1020
|
|
|
Db 961 AAATCTCCAGAAAGAGAAAGAGCAAGAGCGAGCAAAAAGTAAAGTAAAGTAAAGAA 1020
|
|
|
QY 1021 AAGACGAGCAGCTCGAGAGAGCAGAGCGGCAAGAGACAGACAGACAGGCTGATG 1080
|
|
|
Db 1021 AAGACGAGCAGCTCGAGAGAGCAGAGCGGCAAGAGACAGACAGACAGGCTGATG 1080
|
|
|

QY 1081 CAGCAGACTCAGAAAGTGAAGTGAAGTGCCTTTGGAAGACAGATGTTGACCTGAGGACAT 1140
|
|
|
Db 1081 CAGCAGACTCAGAAAGTGAAGTGAAGTGCCTTTGGAAGACAGATGTTGAGACTGAGGACAT 1140
|
|
|
QY 1141 CGTCAGAGAGAAAGTGTGCTCTTTGGCAACGAAAGTGTGTAATGAGAAATGGAACCC 1200
|
|
|
Db 1141 CGTCAGAGAGAAAGTGTGCTCTTTGGCAACGAAAGTGTGTAATGAGAAATGGAACCC 1200
|
|
|
QY 1201 ACCAAGAAATGTTGTCAGAGAGTCCACGTGAGCAACCGTGAAGAAAGACAGAGAGAGAG 1260
|
|
|
Db 1201 ACCAAGAAATGTTGTCAGAGAGTCCACGTGAGCAACCGTGAAGAAAGACAGAGAGAGAG 1260
|
|
|
QY 1261 GAGAGAGAGAGAGAGCTGAAGAGGGGCGTGTGTGAAGAAACAGAGAAATCTTGGCCC 1320
|
|
|
Db 1261 GAGAGAGAGAGAGAGCTGAAGAGGGGCGTGTGTGAAGAAACAGAGAAATCTTGGCCC 1320
|
|
|
QY 1321 CTGAAGAACTGGCTGAGACCCCGAGAGGTCCCGCAGAAAGCTGAGCTGAGAGAGCTGA 1380
|
|
|
Db 1321 CTGAAGAACTGGCTGAGACCCCGAGAGGTCCCGCAGAAAGCTGAGCTGAGAGAGCTGA 1380
|
|
|
QY 1381 TGAAGACAGAGAGATGTGTCTCTGAGAGAGACACACTCAACTGACACACTTAAGTC 1440
|
|
|
Db 1381 TGAAGACAGAGAGATGTGTCTCTGAGAGAGACACACTCAACTGACACACTTAAGTC 1440
|
|
|
QY 1441 CTGAAGAAAGAGCGCTGCCCAACACCCAGAAAGCATTTGTCAGTGAGGTGAGATGCTGT 1500
|
|
|
Db 1441 CTGAAGAAAGAGCGCTGCCCAACACCCAGAAAGCATTTGTCAGTGAGGTGAGATGCTGT 1500
|
|
|
QY 1501 CCTCTCAGAAAGAAATCAAGGTATCAGGAAAGTCCCTTGAAGAACTCTTCAATAGCTCAG 1560
|
|
|
Db 1501 CCTCTCAGAAAGAAATCAAGGTATCAGGAAAGTCCCTTGAAGAACTCTTCAATAGCTCAG 1560
|
|
|
QY 1561 GCTTAAAGAGCTGTCTGGAAGAAAGCAAGAGGGGAAACGAGAGGTGGGGAACCAAG 1620
|
|
|
Db 1561 GCTTAAAGAGCTGTCTGGAAGAAAGCAAGAGGGGAAACGAGAGGTGGGGAACCAAG 1620
|
|
|
QY 1621 AGCTGAGAAATCAACACATTCACACCGAATCCCGAGAGATGATGAGAGAAAG 1680
|
|
|
Db 1621 AGCTGAGAAATCAACACATTCACACCGAATCCCGAGAGATGATGAGAGAAAG 1680
|
|
|
QY 1681 GAGAGAGCTGCGTGTGCTCCCGAGAGAGCTTGAAGAGACCAAGTGTGAGAAAGGGC 1740
|
|
|
Db 1681 GAGAGAGCTGCGTGTGCTCCCGAGAGAGCTTGAAGAGACCAAGTGTGAGAAAGGGC 1740
|
|
|
QY 1741 CGCTGGAAGCAACCCAGAGATGGGGAAGCTGGAAGAGAACTACTTCCATGGAAGAA 1800
|
|
|
Db 1741 CGCTGGAAGCAACCCAGAGATGGGGAAGCTGGAAGAGAACTACTTCCATGGAAGAA 1800
|
|
|
QY 1801 AGAGAGAGGGAGATCACTCCCTGGGATCTTCAAAAAGATGTAACCCCAAGAAACGG 1860
|
|
|
Db 1801 AGAGAGAGGGAGATCACTCCCTGGGATCTTCAAAAAGATGTAACCCCAAGAAACGG 1860
|
|
|
QY 1861 TCCGAAGACCTTCTGAGAGTGAAGAGAGAGAGCTTGAAGAGTCAAGAGCCCACT 1920
|
|
|
Db 1861 TCCGAAGACCTTCTGAGAGTGAAGAGAGAGAGCTTGAAGAGTCAAGAGCCCACT 1920
|
|
|
QY 1921 TGTCTTCACTGATGCAACAGTGTGAAGAAATGCAAGATGAAGTCAAACTGTTGGTGA 1980
|
|
|
Db 1921 TGTCTTCACTGATGCAACAGTGTGAAGAAATGCAAGATGAAGTCAAACTGTTGGTGA 1980
|
|
|
QY 1981 AACAAAGCCAGAGAACCAAGAGCTGAGGTGATATCTTCAAGTCTTGGGAAGCACTGA 2040
|
|
|
Db 1981 AACAAAGCCAGAGAACCAAGAGCTGAGGTGATATCTTCAAGTCTTGGGAAGCACTGA 2040
|
|
|
QY 2041 TTTGTGTGATCATCAAGAGAGAGCAAGAGAGCAATCTTCAATGATGAAGAG 2100
|
|
|
Db 2041 TTTGTGTGATCATCAAGAGAGAGCAAGAGAGCAATCTTCAATGATGAAGAG 2100
|
|
|
QY 2101 GGCCAAAGCACTGGAGAGGGAAGTCAAGAGAGAGAGAGGCTCAGCAAGCAAGAAAG 2160
|
|
|
Db 2101 GGCCAAAGCACTGGAGAGGGAAGTCAAGAGAGAGAGAGGCTCAGCAAGCAAGAAAG 2160
|
|
|
QY 2161 CCGGAACAGACCTGTCTTCTGCAAGCAACCAAGAGAGCAAGAGCAAGAGATTTCT 2220
|
|
|

Db 2161 ||||| CCGGAAACAGAGCTGTTCTTCCAGACACCCAGGAGCAGGACCAAGCCGAAAGGAAGTTCTCT 2220
Qy 2221 CACCCGAGCCAGCGGGAGGCCCTTCCGAGGGGAAGGTGTCTCCACTTTGGAGGTCAITTTA 2280
Db 2221 CACCCGAGCCAGCGGGAGGCCCTTCCGAGGGGAAGGTGTCTCCACTTTGGAGGTCAITTTA 2280
Qy 2281 AAGATTAGTCACTCCCAAGAAAATAATCCAAAGTCAAACTGGGAAGAACCCGAAAGACT 2340
Db 2281 AAGATTAGTCACTCCCAAGAAAATAATCCAAAGTCAAACTGGGAAGAACCCGAAAGACT 2340
Qy 2341 CTAGTGTAGAGCAGTGTCTCCACTGAGATCGAAACCGAGTACGAGGAATCTTTGGGTTTCCA 2400
Db 2341 CTAGTGTAGAGCAGTGTCTCCACTGAGATCGAAACCGAGTACGAGGAATCTTTGGGTTTCCA 2400
Qy 2401 TTAAGAAATTCATCCCGGACCGGGGAAGAAAAGGGCAGACGGGAACCAAGGCCA 2460
Db 2401 TTAAGAAATTCATCCCGGACCGGGGAAGAAAAGGGCAGATGGGAAGCAAGGCCA 2460
Qy 2461 CTGTGGAAGACTCAGGCCAGTGGAGATAATGAGGACGACCTTAATGTCCAGCCGTG 2520
Db 2461 CTGTGGAAGACTCAGGCCAGTGGAGATAATGAGGACGACCTTAATGTCCAGCCGTG 2520
Qy 2521 TGCCCTGTCTGAGTATAATGCAAGTGGAGGAGAGATGGGAAGCCACGGGAATACGG 2580
Db 2521 TGCCCTGTCTGAGTATGATGAGTGGAGGAGAGATGGGAAGCCACGGGAATACGG 2580
Qy 2581 AGTGCCCCAGCTGCTGGGGGCTGTGTACGTGTCCGAGGAGCTCAGTAAGACTCTGGTCC 2640
Db 2581 AGTGCCCCAGCTGCTGGGGGCTGTGTACGTGTCCGAGGAGCTCAGTAAGACTCTGGTCC 2640
Qy 2641 ACAGTGTAGTGTGCGAGTCAATGATGGGACCGAGGCACTGTCAGTGTGCGAAGCGGT 2700
Db 2641 ACAGTGTAGTGTGCGAGTCAATGATGGGACCGAGGCACTGTCAGTGTGCGAAGCGGT 2700
Qy 2701 CTCTTCGTGATATCCGTTCCGTACAGAACCTTTGTAACACACAGCGGGAGAACCCA 2760
Db 2701 CTCTTCGTGATATCCGTTCCGTACAGAACCTTTGTAACACACAGCGGGAGAACCCA 2760
Qy 2761 TGCCACCTGTTGAAGAGGTCACTGAAAGACATCATTTGAGAGAAACTCCTGTGTCTCA 2820
Db 2761 TGCCACCTGTTGAAGAGGTCACTGAAAGACATCATTTGAGAGAAACTCCTGTGTCTCA 2820
Qy 2821 CCAGAGCTTACAGAGGTAAAGATGCCATGACGACATGTCACAGTGAAGTGAAT 2880
Db 2821 CCAGAGCTTACAGAGGTAAAGATGCCATGACGACATGTCACAGTGAAGTGAAT 2880
Qy 2881 TCACCTCAGAACTGTGACAGCCACAGAGACCTCAGAGGCTCTCCGTACTGAAAGAGTTA 2940
Db 2881 TCACCTCAGAACTGTGACAGCCACAGAGACCTCAGAGGCTCTCCGTACTGAAAGAGTTA 2940
Qy 2941 CCGAAGCATCGGGGGCCGAGAGACCAACAGATGTTGTCGAGATTTCCAGCTGACTG 3000
Db 2941 CCGAAGCATCGGGGGCCGAGAGACCAACAGATGTTGTCGAGATTTCCAGCTGACTG 3000
Qy 3001 ACTCCGAGACACACAGAGGAAGCCACCCAGTTACGAGGTAGAGAGTGTGTCTAG 3060
Db 3001 ACTCCGAGACACACAGAGGAAGCCACCCAGTTACGAGGTAGAGAGTGTGTCTAG 3060
Qy 3061 ATACAGAAAGAGGAGCGCCAGACCGAGGCCATCTCCAGCGGTTGACACAGGTGA 3120
Db 3061 ATACAGAAAGAGGAGCGCCAGACCGAGGCCATCTCCAGCGGTTGACACAGGTGA 3120
Qy 3121 AAGAGGAGTCCCAGGTGCTTGCACACCCAGACTGTGACAGAAACGGGGTCAAAAGCACTGG 3180
Db 3121 AAGAGGAGTCCCAGGTGCTTGCACACCCAGACTGTGACAGAAACGGGGTCAAAAGCACTGG 3180
Qy 3181 AGAAGGTTAGAGGAGTGAAGAGGACTCCGAAGTGTGCTTCCGAGAAAGAGAGGACG 3240
Db 3181 AGAAGGTTAGAGGAGTGAAGAGGACTCCGAAGTGTGCTTCCGAGAAAGAGAGGACG 3240
Qy 3241 TTATGCCAAAGAACCCGTCGAGGAAGCTTGAGCTATCTTGACAGGGCTCTGAGA 3300

Db 3241 TTATGCCAAAGAACCCGTCGAGGAAGCTGGAGCTGAGCATCTTGCACAGGGCTCTGAGA 3300
Qy 3301 CTGGACAGGCTACTCCAGAGAGCCTTTGAAGTTCCTGAAGTCAACGGCAGATGTAGACCATG 3360
Db 3301 CTGGACAGGCTACTCCAGAGAGCCTTTGAAGTTCCTGAAGTCAACGGCAGATGTAGACCATG 3360
Qy 3361 TCGCCACGTGCGCAGGTTATCAAGCTCCAGCAGCTGATGGAAACAGGCCGTGGCCCTTGAGT 3420
Db 3361 TCGCCACGTGCGCAGGTTATCAAGCTCCAGCAGCTGATGGAAACAGGCCGTGGCCCTTGAGT 3420
Qy 3421 CATCCGAAACCTTTGACAGCAGTGAGACAAATGGAAGCACTCCTCTAGCAGATTCAGACA 3480
Db 3421 CATCCGAAACCTTTGACAGCAGTGAGACAAATGGAAGCACTCCTCTAGCAGATTCAGACA 3480
Qy 3481 CTCAGATGGGACACAGCAAGATGAAACCAATTCACGCCAGGACAGTAAAGCCACTGTCAG 3540
Db 3481 CTCAGATGGGACACAGCAAGATGAAACCAATTCACGCCAGGACAGTAAAGCCACTGTCAG 3540
Qy 3541 CTGTCAAGCAGTCAAGGTCACAGAGAGAGAGGGCGCTACTGCTCAGAAAGAGAGGCCCTT 3600
Db 3541 CTGTCAAGCAGTCAAGGTCACAGAGAGAGAGGGCGCTACTGCTCAGAAAGAGAGGCCCTT 3600
Qy 3601 CGACACTTACCTAATAATGTTCCAGCCAGGAAGAAATGCGGGAAGAACCCAGGAAGAGATG 3660
Db 3601 CGACACTTACCTAATAATGTTCCAGCCAGGAAGAAATGCGGGAAGAACCCAGGAAGAGATG 3660
Qy 3661 TTCTTGAACCTTACACAGCAAGAGCTTACTGCTCAGCCGCTGCCCGTTCTGGCAAAAGACTG 3720
Db 3661 TTCTTGAACCTTACACAGCAAGAGCTTACTGCTCAGCCGCTGCCCGTTCTGGCAAAAGACTG 3720
Qy 3721 AGTGCGGTCAAGAGGTGAGGTTGACTGCTGATGGGAAAAAAGTCAAAAGAAAGACAGG 3780
Db 3721 AGTGCGGTCAAGAGGTGAGGTTGACTGCTGATGGGAAAAAAGTCAAAAGAAAGACAGG 3780
Qy 3781 AGTGTTTGTACACTCTGGACCCAAACAGTCAAAAGCTCTGATGTGACATATGACAGTG 3840
Db 3781 AGTGTTTGTACACTCTGGACCCAAACAGTCAAAAGCTCTGATGTGACATATGACAGTG 3840
Qy 3841 AAGTGATGGGAGTGGCGGGGTGTGAGAAAAAGGAGAGTACTGAAAGTGCAGAGCTTTAGCC 3900
Db 3841 AAGTGATGGGAGTGGCGGGGTGTGAGAAAAAGGAGAGTACTGAAAGTGCAGAGCTTTAGCC 3900
Qy 3901 TGAGAGGAGGAGATGGAATACTGACGTTGAAAGAGAAAAAGGAGACAAAGCCAGAGC 3960
Db 3901 TGAGAGGAGGAGATGGAATACTGACGTTGAAAGAGAAAAAGGAGACAAAGCCAGAGC 3960
Qy 3961 AAGTGAGTGAAGAGGTGAGCAGGAACAGCCGCTCTGAGCATGAAGAACCTTACGGGA 4020
Db 3961 AAGTGAGTGAAGAGGTGAGCAGGAACAGCCGCTCTGAGCATGAAGAACCTTACGGGA 4020
Qy 4021 AGCCAGTCTCTGACACTTGACATGCCAGCTCAGAGAGGGGGAAGGCACTGGGAAGCCTTG 4080
Db 4021 AGCCAGTCTCTGACACTTGACATGCCAGCTCAGAGAGGGGGAAGGCACTGGGAAGCCTTG 4080
Qy 4081 GAGGAAGCCCTTCTCTCCAGACCAAGACAAAGCAGGTTGCATAGAGGTTCAAGTTCAAA 4140
Db 4081 GAGGAAGCCCTTCTCTCCAGACCAAGACAAAGCAGGTTGCATAGAGGTTCAAGTTCAAA 4140
Qy 4141 GCCTGACACAAACAGTCACTCAAAACAGCAGAGCTGTGGAAGAGGTCATGAAGACGGTTG 4200
Db 4141 GCCTGACACAAACAGTCACTCAAAACAGCAGAGCTGTGGAAGAGGTCATGAAGACGGTTG 4200
Qy 4201 TGATTTTCAGAGCAGGTGAAAGTCCAGAGTGTGTAGTGCACACTTTATTACAGCTGAGA 4260
Db 4201 TGATTTTCAGAGCAGGTGAAAGTCCAGAGTGTGTAGTGCACACTTTATTACAGCTGAGA 4260
Qy 4261 AGTCTCTTCCAAACGGGTGGCCACTGGACTCTTTAGCATGACAGGACACGGGTACCCCTGG 4320
Db 4261 AGTCTCTTCCAAACGGGTGGCCACTGGACTCTTTAGCATGACAGGACACGGGTACCCCTGG 4320
Qy 4321 GGCCTGAGTCTCAGGAGAGTCCATCCCAATCATAGTAATCTCTGCTCTCTGTAAGACACC 4380
Db 4321 GGCCTGAGTCTCAGGAGAGTCCATCCCAATCATAGTAATCTCTGCTCTCTGTAAGACACC 4380


```

Qy 4381 TACATCTGACCTTACAGAGAAATTAAGCGATCCACAGAGAGCTGATCAAGAGAAAGG 4440
Db 4381 TACATCTGACCTTACAGAGAAATTAAGCGATCCACAGAGAGAGATCAAGAGAAAGG 4440
Qy 4441 ACAAGCCAGATGCTGATCTGATGCTGACGGCAAGAGAGATACAGCATCGAAAAAGTCC 4500
Db 4441 ACAAGCCAGATGCTGATCTGATGCTGACGGCAAGAGAGATACAGCATCGAAAAAGTCC 4500
Qy 4501 TCAAGGCTGAAGCTGAGATCTCTGAACTTGAAGATGAAGCAACAAGATTGTCTGAACG 4560
Db 4501 TCAAGGCTGAAGCTGAGATCTCTGAACTTGAAGATGAAGCAACAAGATTGTCTGAACG 4560
Qy 4561 TCATTCAGACAGCCGTTTACAGATTGCACTGACAGAAACAGCCCCCGAAATCTCATGCTT 4620
Db 4561 TCATTCAGACAGCCGTTTACAGATTGCACTGACAGAAACAGCCCCCGAAATCTCATGCTT 4620
Qy 4621 ATGATTCACAGACCCAGGTTCTGATGCAAGCTTGACAGAGAGAGCCCAACAGATGCT 4680
Db 4621 ATGATTCACAGACCCAGGTTCTGATGCAAGCTTGACAGAGAGAGCCCAACAGATGCT 4680
Qy 4681 GGACAAATGAAGATGCTCAAGATGAACAACCCAGTCCGAGCCAGAGAGAGAGAGAGAGAG 4740
Db 4681 GGACAAATGAAGATGCTCAAGATGAACAACCCAGTCCGAGCCAGAGAGAGAGAGAGAG 4740
Qy 4741 AAGTCCTGACCGTTCTGAGAGGATGGGCTCAGCTCTGAAATGCTTGCCGCTTGCACT 4800
Db 4741 AAGTCCTGACCGTTCTGAGAGGATGGGCTCAGCTCTGAAATGCTTGCCGCTTGCACT 4800
Qy 4801 TGAAGGCGCCGTTCTCAAGATGAAGATGAAGAGTCTCTCAACCAAGATTCGAA 4860
Db 4801 TGAAGGCGCCGTTCTCAAGATGAAGATGAAGAGTCTCTCAACCAAGATTCGAA 4860
Qy 4861 AGAGCATGCTGCTGATGAGCTCTCAGCTCAAGCTTGAAGCTGAGCCAGAGAGAGAGAGCT 4920
Db 4861 AGAGCATGCTGCTGATGAGCTCTCAGCTCAAGCTTGAAGCTGAGCCAGAGAGAGAGAGCT 4920
Qy 4921 CTGGAACCTTAACAAAGATCCCAAGATCCCAAGATCCCAAGATCCCAAGATCCCAAGAT 4980
Db 4921 CTGGAACCTTAACAAAGATCCCAAGATCCCAAGATCCCAAGATCCCAAGATCCCAAGAT 4980
Qy 4981 ATCCCAAGATCCCAAGATCCCAAGATCCCAAGATCCCAAGATCCCAAGATCCCAAGAT 5040
Db 4981 ATCCCAAGATCCCAAGATCCCAAGATCCCAAGATCCCAAGATCCCAAGATCCCAAGAT 5040
Qy 5041 AGGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5100
Db 5041 AGGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5100
Qy 5101 TTAGTGTCTCATTTGATCATCTGTGAAGACCAAGATGTGAACCAAGTCAAGAACCAAGATG 5160
Db 5101 TTAGTGTCTCATTTGATCATCTGTGAAGACCAAGATGTGAACCAAGTCAAGAACCAAGATG 5160
Qy 5161 CTGCTGTTGGAGCTTTGAGACCAAGATTTCAAGGCCATGATCCAGAGAGAGAGAGAGAG 5220
Db 5161 CTGCTGTTGGAGCTTTGAGACCAAGATTTCAAGGCCATGATCCAGAGAGAGAGAGAGAG 5220
Qy 5221 TCCAGTGAATTTCCACC 5237
Db 5221 TCCAGTGAATTTCCACC 5237
Qy 5220 TCCAGTGAATTTCCACC 5236
Db 5220 TCCAGTGAATTTCCACC 5236

```

```

RESULT 3
LOCUS AR071282 5134 bp DNA linear PAT 18-FEB-2000
DEFINITION Sequence 1 from patent US 5910442.
ACCESSION AR071282
VERSION AR071282.1 GI:7222170
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 5134)

```

```

AUTHORS Gelman, I.H.
TITLE Tumor suppressor gene
JOURNAL Patent: US 5910442-A 1 08-JUN-1999;
FEATURES
Source Location/Qualifiers
1..5134
/organism="unknown"
BASE COUNT 1511 a 1167 c 1451 g 1005 t
ORIGIN
Query Match 78.4%; Score 4831.4; DB 6; Length 5134;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 5049; Conservative 0; Mismatches 76; Indels 31; Gaps 14;
Qy 1016 GGAAGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1075
Db 1 GGAAGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
Qy 1076 GTTGTCAGACACTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1135
Db 61 GTTGTCAGACACTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
Qy 1136 GGATTCGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1195
Db 121 GGATTCGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
Qy 1196 AGCCCAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1255
Db 181 AGCCCAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
Qy 1256 GCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1315
Db 241 GCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
Qy 1316 GCCCCCTGAGAACTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1375
Db 301 GCCCCCTGAGAACTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
Qy 1376 GCTGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1435
Db 361 GCTGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Qy 1436 AAGTCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1495
Db 421 AAGTCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
Qy 1496 GCTGCTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1555
Db 481 GCTGCTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Qy 1556 CTCAGAGCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1615
Db 541 CTCAGAGCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Qy 1616 CGAAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1675
Db 601 CGAAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Qy 1676 GAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1735
Db 661 GAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Qy 1736 AGGGCCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1795
Db 721 AGGGCCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 775
Qy 1796 GAAGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1855
Db 776 GAAGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 835
Qy 1856 ACAGGTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1915
Db 836 AC-GGTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 894

```

QY 1916 CACCTTGTCTCTCCTCCTATAGCACAGTGTGAGAAATGCAAGATCAAGTCAAACTGTTGG 1975
Db 895 CACCTTGTCTCTCCTCCTATAGCACAGTGTGAGAAATGCAAGATCAAGTCAAACTGTTGG 954
QY 1976 TGAGGAAACAAAGCCAGAGGAAACCAAGCGTAGGGTGGATCTTCACTGTCTTGGGAAGC 2035
Db 955 TGAGGAAACAAAGCCAGAGGAAACCAAGCGTAGGGTGGATCTTCACTGTCTTGGGAAGC 1014
QY 2036 ACTGATTTTGTGTGGATCATCTCAAGAAAGAGAGAGCAAGGAGCATCTCTTCTCAGATGATGA 2095
Db 1015 ACTGATTTTGTGTGGATCATCTCAAGAAAGAGAGCAAGGAGCATCTCTTCTCA--GATAT 1071
QY 2096 AGGAGGCCAAGAGACACTGGGA--GGGACAGTCAACAGAGAGAGGAGGCCAGCAAGACACA 2154
Db 1072 AAGAGGCCAAGAGACACTGGGAGGGGAGACAGTCAACAGAGAGAGGAGGCCAGCAAGACACA 1131
QY 2155 AAGAAGCCGGAACAGACAGCTGTCTTCTCCAGACACCCAGGAGCAGGACCAAGCGCAAGGAA 2214
Db 1132 AAGAAGCC--GAACAGACGCTGTCTTCTCCAGACACCCAGGAGCAGGACCAAGCGCAAGGAA 1190
QY 2215 GTTCTCTACCCGAGCCAGCGGGAAGCCCTTCCGAAGGGGAAGGTGTCTTCCACTTTGGGAGT 2274
Db 1191 GTTCTCTACCCGAGCCAGCGGGAAGCCCTTCCGAAGGGGAAGGTGTCTTCCACTTTGGGAGT 1250
QY 2275 CATTTAAAGATTAGTCACTCCCAAGAAABAATCCAAAGTCAAACTGGAAGAGAAAGCCG 2334
Db 1251 CATTTAAAGATTAGTCACTCCCAAGAAABAATCCAAAGTCAAACTGGAAGAGAAAG 1310
QY 2335 AAGACTCTAGTGTAG-----AGCAGTTGTCCACTGAGATCGAAACCGAGTAGAGAAGA 2386
Db 1311 CCGGAAGGACTCTAGTTGTAGGAGCAGGTTGTCCACTGAGATCGAACCGTGTAGAGNAGA 1370
QY 2387 ATCTTGGGTTTCCATTAAAGAAATTCATPCCCGGACGCGGGAAGAAAGGCGACAGCGGAA 2446
Db 1371 ATCTTGGGTTTCCATTAAAGAAATTCATPCCCGGACGCGGGAAGAAAGGCGACATGGAA 1430
QY 2447 -GCAAGAACCAAGCCACTGTGGAAGACTCAGGCGCAGTGGAGATAAATGAGGACGACCTTA 2505
Db 1431 GGCAAGAACCAAGCCACTGTGGAAGACTCAGGCGCAGTGGAGATAAATGAGGAGCGGCTG 1490
QY 2506 ATGTCCAGCGCTCGTGCTGTCTGAGTATAATGAGTGGAGAGGGAAGATGGAAG 2565
Db 1491 ATGTCCAGCGCTCGTGCTGTCTGAGTATGATGAGTGGAGAGGGAAGATGGAAG 1550
QY 2566 CCAGGGGAATACGGAGCTGCCAGCTGCCAGCTGTCTGGGCTGTGT--AGTCTCCGAGGAGTCA 2625
Db 1551 CCAGGGGAATGCGGAGCTGCCAGCTGCCAGCTGTCTGGGCTGTGT--AGTCTCCGAGGAGTCA 1607
QY 2626 GTAAGACTCTGGTCCACACTGTGAGTGTCCAGTCAATTGATGGGACGAGGCGAGTCACCA 2685
Db 1608 GTAAGACTCTGGTCCACACTGTGAGTGTCCAGTCAATTGATGGGACGAGGCGAGTCACCA 1667
QY 2686 GTGTGGAAGCGGTCTCTCTTCTGGATATCCGCTTCCGTAACAGAACCTCTTGAACACA 2745
Db 1668 GTGTGGAAGCGGTCTCTCTTCTGGATATCCGCTTCCGTAACAGAACCTCTTGAACACA 1727
QY 2746 CAGCGGAGAACCCATGCCACTGTTGAAGAGTCACTGAAAGAGACATCAITTCGAGAAG 2805
Db 1728 CAGCGGAGAACCCATGCCACTGTTGAAGAGTCACTGAAAGAGACATCAITTCGAGAAG 1787
QY 2806 AACTCTGTGCTCACCAGACCTTACCAGAGGTTAAGATGCCCATGAGCATGGTCA 2865
Db 1788 AACTCTGTGCTCACCAGACCTTACCAGAGGTTAAGATGCCCATGAGCATGGTCA 1847
QY 2866 CCAGTGAAGTGGATTTTCACTCTCAGAAGCTGTGACAGCCACAGAGACTCAGAGGCTCTCC 2925
Db 1848 CCAGTGAAGTGGATTTTCACTCTCAGAAGCTGTGACAGCCACAGAGACTCAGAGGCTCTCC 1907
QY 2926 GTACTGAAGAAGTTACCGAAGCATTCGGGGCCGAAGAGACCAACAGACATGGTGTCCGCGAG 2985
Db 1908 GTACTGAAGAAGTTACCGAAGCATTCGGGGCCGAAGAGACCAACAGACATGGTGTCCGCGAG 1967
QY 2986 TTTCCAGCTGACTGACTCCCCAGACACACACAGAGGAGCCACCCAGCTTCAGGAGGTAG 3045

Db 1968 TTTCCAGCTGACTGACTCCCCAGACACACAGAGAAAGCCACCCAGTTTCAGGAGGTAG 2027
QY 3046 AGAGTGTGTCTGTAGATACAGAAAGAGAGGCGCCAGACGCGAGGCGATCTCTCCAAGCCG 3105
Db 2028 AGGCTGTGTCTGTAGATACAGAAAGAGAGGCGCCAGACGCGAGGCGATCTCTCCAAGCCG 2087
QY 3106 TTGCACACAAGGTGAAAGAGGAGTCCCAGGTGCTCTCAACCCAGACTGTGCAGAGAACGG 3165
Db 2088 TTGCACACAAGGTGAAAGAGGAGTCCCAGGTGCTCTCAACCCAGACTGTGCAGAGAACGG 2147
QY 3166 GGTCAAAAGCACTGGAGAAAGTTTGAAGAGGTAGAGAGGACTCCGAAGTGTGCTGGCTTCGG 3225
Db 2148 GGTCAAAAGCACTGGAGAAAGTTTGAAGAGGTAGAGAGGACTCCGAAGTGTGCTGGCTTCGG 2207
QY 3226 AGAAAGAGAAAGGACGTTATGTCGCAAAAGGACCCGTGCAGGAAGCTGTGAGCTGAGCATCTTG 3285
Db 2208 AGAAAGAGAAAGGACGTTATGTCGCAAAAGGACCCGTGCAGGAAGCTGTGAGCTGAGCATCTTG 2267
QY 3286 CACAGGGCTCTGAGACTGTGACAGGGCTATCTCAGAGAGCCCTTGAAGTTCCTGAAAGTCAAG 3345
Db 2268 CACAGGGCTCTGAGACTGTGACAGGGCTACTCCAGAGAGCCCTTGAAGTTCCTGAAAGTCAAG 2327
QY 3346 CAGATGTAGACCATGTCCGACAGCTGCGCAGGTTATCAAGCTCCAGAGCTGATGGAACAGG 3405
Db 2328 CAGATGTAGACCATGTCCGACAGCTGCGCAGGTTATCAAGCTCCAGAGCTGATGGAACAGG 2387
QY 3406 CCGTGCCCTCTGAGTCACTCCGAAACCTTGACAGACAGTGTGAGCAAAATGGAAGCACTCCCT 3465
Db 2388 CCGTGCCCTCTGAGTCACTCCGAAACCTTGACAGACAGTGTGAGCAAAATGGAAGCACTCCCT 2447
QY 3466 TAGCAGATTTCAGACACTGTGAGATGGGACACAGCAAGATGAAACCATTTGACAGCAGGACA 3525
Db 2448 TAGCAGATTTCAGACACTGTGAGATGGGACACAGCAAGATGAAACCATTTGACAGCAGGACA 2507
QY 3526 GTAAAGCCACTGCGAGCTGTCCAGCAGTCAAGGTTCACAGAAAGAGGCGGCTACTGTCTC 3585
Db 2508 GTAAAGCCACTGCGAGCTGTCCAGCAGTCAAGGTTCACAGAAAGAGGCGGCTACTGTCTC 2567
QY 3586 AGAAAGAGGAGCTTCCACACTACCTTAATGTTTCCAGCCAGGAAGAACATGGGGAAG 3645
Db 2568 AGAAAGAGGAGCTTCCACACTACCTTAATGTTTCCAGCCAGGAAGAACATGGGGAAG 2627
QY 3646 AACCCAGAAAGAGATGTTCTTGAACCTTACACAGCAAGAGCTTGTCTGCGAGCGCTGCCCG 3705
Db 2628 AACCCAGAAAGAGATGTTCTTGAACCTTACACAGCAAGAGCTTGTCTGCGAGCGCTGCCCG 2687
QY 3706 TTCTGCAAAAGACTGAGGTGGGTCAAGAGGTGAGGTGACTGGTTGGATGGAGAAAAAG 3765
Db 2688 TTCTGCAAAAGACTGAGGTGGGTCAAGAGGTGAGGTGACTGGTTGGATGGAGAAAAAG 2747
QY 3766 TCAAGAGAAACAGGAGGTGTTGTACACTCTGGACCCCAACAGTCAAAAGGCTGTGTGATG 3825
Db 2748 TCAAGAGAAACAGGAGGTGTTGTACACTCTGGACCCCAACAGTCAAAAGGCTGTGTGATG 2807
QY 3826 TGACATATGACAGTGAAGTGTGGAGTGGCGGGTGTGAGGAAAGAGAGACTTCTGAAG 3885
Db 2808 TGACATATGACAGTGAAGTGTGGAGTGGCGGGTGTGAGGAAAGAGAGACTTCTGAAG 2867
QY 3886 TGCAGAGTCTTACCTTGGAGGAGGAGATGGAATACTGACGTTGAAAGGAGAAAAAGGG 3945
Db 2868 TGCAGAGTCTTACCTTGGAGGAGGAGATGGAATACTGACGTTGAAAGGAGAAAAAGGG 2927
QY 3946 AGACAAAGCCAGACAGTGAAGTGAAGAGGTGAGCAGGAAACAGCCGCTCTCTGAGCATG 4005
Db 2928 AGACAAAGCCAGACAGTGAAGTGAAGAGGTGAGCAGGAAACAGCCGCTCTCTGAGCATG 2987
QY 4006 AAGGAACCTTACCGGGAAGCCAGTCTTGACATTTGACATGCCCCAGCTCAGAGAGGGGAAGG 4065
Db 2988 AAGGAACCTTACCGGGAAGCCAGTCTTGACACTTGAATGCCAGCTCAGAGAGGGGAAGG 3047
QY 4066 CACTGGGAAGCCTTGGAGGAAGCCCTTCTCTCCAGACCAAGCAAGCAGGTTGATGATG 4125

Db 3048 CACTGGGAAGCCTTGGAGGAAGCCCTTCTCTCCAGACCAAGCAAGAGGTTGCATAG 3107
Qy 4126 AGGTTCAAGTTCAAAAGCCTTGGAGCAACAAGTCACTCAAAAGAGAGAGCTGTGAAAAAG 4185
Db 3108 AGGTTCAAGTTCAAAAGCCTTGGAGCAACAAGTCACTCAAAAGAGAGAGCTGTGAAAAAG 3167
Qy 4186 TCATAGAAACGGTGTGTGATTTTCAGAGACAGGTGAAAGTCCAGAGTGTGTAGTGCACAT 4245
Db 3168 TCATAGAAACGGTGTGTGATTTTCAGAGACAGGTGAAAGTCCAGAGTGTGTAGTGCACAT 3227
Qy 4246 TATTAACCAAGCTGAGAAAGTCTCTGCAACGGGTGGCCACTGAGACTTTAGCATGAGAGG 4305
Db 3228 TATTAACCAAGCTGAGAAAGTCTCTGCAACGGGTGGCCACTGAGACTTTAGCATGAGAGG 3287
Qy 4306 ACACGGTACCCCTGGGGGCTGAGTCTCAGGCAAGAAATCCATCCATCATAGTAATCTCTG 4365
Db 3288 ACAAGGTACCCCTGGGGGCTGAGTCTCAGGCAAGAAATCCATCCATCATAGTAATCTCTG 3347
Qy 4366 CTCTGAAAGCACTTACATCTGTACCTAACAAGGAAATTAAGCCGATCCCAAGAGAGGC 4425
Db 3348 CTCTGAAAGCACTTACATCTGTACCTAACAAGGAAATTAAGCCGATCCCAAGAGAGGC 3407
Qy 4426 GATGAGAAAGAGAGCAAGCCAGATGCTGTGCTGATGCTGAGCGCAAGAGAGATACAG 4485
Db 3408 GATGAGAGAGAGAGCAAGCCAGATGCTGTGCTGATGCTGAGCGCAAGAGAGATACAG 3467
Qy 4486 CAATGAAAAAGTCCCTCAAGGCTGAACCTGAGATCTGTGAACTTGAAGTAAGCAACA 4545
Db 3468 CAATGAAAAAGTCCCTCAAGGCTGAACCTGAGATCTGTGAACTTGAAGTAAGCAACA 3527
Qy 4546 AGATTGTGCTGAACGTCAATTCAGACAGCCGTGACAGTTCGCCAGCTACAGAAACAGCCC 4605
Db 3528 AGATTGTGCTGAACGTCAATTCAGACAGCCGTGACAGTTCGCCAGCTACAGAAACAGCCC 3587
Qy 4606 CGGAAATCTAGCTTATGATTTACAAGACCCAGATTCCTGATGAGAGGCTTGAACAGAGG 4665
Db 3588 CGGAAATCTAGCTTATGATTTACAAGACCCAGATTCCTGATGAGAGGCTTGAACAGAGG 3647
Qy 4666 AGCCCAACAGATGCTGACAAATAATGAAGATGCCAAGATGAACAACCCAGTCCGACAGC 4725
Db 3648 AGCCCAACAGATGCTGACAAATAATGAAGATGCCAAGATGAACAACCCAGTCCGACAGC 3707
Qy 4726 CCAGAGAGGACTTGCAGATCTGACCCGTTCTGAGGCAATGGGCTCAGCCTCGGAAATGCT 4785
Db 3708 CCAGAGAGGACTTGCAGATCTGACCCGTTCTGAGGCAATGGGCTCAG- -CTCGGAAATGCT 3765
Qy 4786 TGGCGGCTTTCAGATTTGAAGGCGCGGTGCAAGTAAGATTTGAAGAGTGGCTCTCTCA 4845
Db 3766 TGGCGGCTTTCAGATTTGAAGGCGCGGTGCAAGTAAGATTTGAAGAGTGGCTCTCTCA 3825
Qy 4846 ACCCAAGATCCAAAAGAGAGATGCTGATGGCCCTCAGCTCCAAAGCTTAAGCCAGG 4905
Db 3826 ACCCAAGAT - CAAAAGAGAGATGCTGATGGCCCTCAGCTCCAAAGCTTAAGCCAGG 3884
Qy 4906 CAGAGGCCAGTGCCTGTGAAACCTTAACCAAGATTCGCCAGACCAACAGGACCAAGC 4965
Db 3885 CAGAGGCCAGTGCCTGTGAAACCTTAACCAAGATTCGCCAGACCAACAGGACCAAGC 3941
Qy 4966 TAAACGAGAGAGGCGCATCCCCCAAAAGTTGAGGTCCAGGAAGAAATGTACCAAGT 5025
Db 3942 TAAACGAGAGAGG - ATGCCCCCAAAAGTTGAGGTCCAGGAAGAAATGTACCAAGT 4000
Qy 5026 CAGTCAAGAGAGAGGCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5085
Db 4001 CAGTCAAGAGAGAGAGGCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4060
Qy 5086 CAGAAATCTAAGATGTAGTGTCTCATTTGTACATCTGTAAAGACAGATGTGAAAAAG 5145
Db 4061 CAGAAATCTAAGATGTAGTGTCTCATTTGTACATCTGTAAAGACAGATGTGAAAAAG 4120
Qy 5146 TCACAGAGAGAGATGCTGCTGTGTTGGAGAGCTTTGAGACCAATTTTCAGAGCCCATGAGATC 5205
Db 4121 TCACAGAGAGAGATGCTGCTGTGTTGGAGAGCTTTG - GACCAAGATTTTCAGAGCCCATGAGATC 4179

Qy 5206 CAGAGAGAGAGGCGGTCGAATGATTTCCACACAGTAGAGAGAGAGAGAGAGAGAGAGAG 5265
Db 4180 CAGAGAGAGAGGCGGTCGAATGATTTCCACACAGTAGAGAGAGAGAGAGAGAGAGAGAG 4239
Qy 5266 TCATGGAGAGTAG 5325
Db 4240 TCATGGAGAGTAG 4299
Qy 5326 TTGATGCCGTCGTTGATTTCTTAACATTTCTGCTGTTCAAGAGCTGCTTTGATTTGCCCC 5385
Db 4300 TTGATGCCGTCGTTGATTTCTTAACATTTCTGCTGTTCAAGAGCTGCTTTGATTTGCCCC 4359
Qy 5386 TTGATGCCGTCGTTGATTTCTTAAGGTTAAGGTCCTGAGTCTCAACCTGGAACCAATTC 5445
Db 4360 TTGATGCCGTCGTTGATTTCTTAAGGTTAAGGTCCTGAGTCTCAACCTGGAACCAATTC 4419
Qy 5446 GCCATACCTAGTTCACATTCCTCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5505
Db 4420 GCCATACCTAGTTCACATTCCTCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4479
Qy 5506 TTTATGATGCTCTCTCTCTGTAATTTGATTTATTTTCTTAACGTTTAAGCAATGCT 5565
Db 4480 TTTATGATGCTCTCTCTCTGTAATTTGATTTATTTTCTTAACGTTTAAGCAATGCT 4539
Qy 5566 TTTTGTATTAAGCAATATTAAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5625
Db 4540 TTTTGTATTAAGCAATATTAAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4599
Qy 5626 GCCTCAACTGATACCTGAG 5685
Db 4600 GCCTCAACTGATACCTGAG 4659
Qy 5686 TTAAGTTTACTGATGCTTAAGTCTGTGGGCTCTAGTCCCTGAAATGATGTTGTTCC 5745
Db 4660 TTAAGTTTACTGATGCTTAAGTCTGTGGGCTCTAGTCCCTGAAATGATGTTGTTCC 4719
Qy 5746 TATGACAG 5805
Db 4720 TATGACAG 4779
Qy 5806 TACATGATTTTTCCTCCCTTTTGTATATCAAGTCCAGGTTGAAAGAGTCTCTCT 5865
Db 4780 TACATGATTTTTCCTCCCTTTTGTATATCAAGTCCAGGTTGAAAGAGTCTCTCTCT 4839
Qy 5866 GTGTGATTAAGCCCTGCTCTTATGATATGACAAATGAGTGTGCTTAAGGCCATGA 5925
Db 4840 GTGTGATTAAGCCCTGCTCTTATGATATGACAAATGAGTGTGCTTAAGGCCATGA 4899
Qy 5926 GATGTTTCTTAATGAG 5985
Db 4900 GATGTTTCTTAATGAG 4959
Qy 5986 GGACCGAATTCATATGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6045
Db 4960 GGACCGAATTCATATGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5019
Qy 6046 AATCTGAGTGTGTGTGTATATCTGTGTCCTCTTTTAAGAACATGTTCATATG 6105
Db 5020 AATCTGAGTGTGTGTGTATATCTGTGTCCTCTTTTAAGAACATGTTCATATG 5079
Qy 6106 TTCTTTTGAATTAATTTGATTTGAACAATGATTTGAATTAATTAATTAATTTGACTAC 6160
Db 5080 TTCTTTTGAATTAATTTGATTTGAACAATGATTTGAATTAATTAATTAATTTGACTAC 5134

RESULT 4
AF326228
LOCUS AF326228 6197 bp mRNA linear ROD 16-JUL-2001
DEFINITION Mus musculus PKC binding protein SseCKS mRNA, complete cds.
ACCESSION AF326228
VERSION AF326228.1 GI:13236403
KEYWORDS

SOURCE Mus musculus.
ORGANISM Mus musculus.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Camus, A., Mesbah, K., Rallu, M., Babinet, C. and Barra, J.
TITLE Gene trap insertion reveals two open reading frames in the mouse SseCKs gene: the form predominantly detected in the nervous system is suppressed by the insertion while the other, specific of the testis, remains expressed
JOURNAL Mech. Dev. 105 (1-2), 79-91 (2001)
MEDLINE 21322683
PUBMED 11429284
REFERENCE 2 (bases 1 to 6197)
AUTHORS Camus, A., Mesbah, K., Babinet, C. and Barra, J.
TITLE Direct Submision
JOURNAL Submitted (06-DEC-2000) Immunologie, Institut Pasteur, 25 rue du Docteur Roux, Paris 75015, France
FEATURES Location/Qualifiers
source 1..6197
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="10"
140..5194
/codon_start=1
/product="PKC binding protein SseCKs"
/protein_id="AAK16150.1"
/db_xref="GI:13236404"
/translation="MGAGSTSEPOEASDTPSELHSGHPAAEASGAAGDPAD
ADPATKLPOKQSLAVNGVAREQDVHVOESQDOGEFVTVEDVGOESEDVKEKDR
AKEMAASSTVEDITKDEQETPEIIEQIPASESNVEAQAESQANDGVKKVFKF
VGFKFTVKDKNEKSDTVOLLTKDEGEGEASVAGDHQBPQVETVGESAKSESEL
KOSTEQEETLKQASSTPEIQAESQGGTTEEAQKQDEENREKPTKLESPTPSP
NETTSFFKFTFWAGWRKTSFKFKEDDLLETSKREKEAEKVDEBEGSKTPAP
ABEQEPAEGTDOARLSADYKEVLEPLEDOVGLSEKCAPLATEVDEKTEAHOEV
VAEVHTVETKTKGGAEGVSDVVEGSESLPEKLAETOEVPQEPVPEALMKT
KEVCVSGDHTQLTDLSPREKMLPKHPEGIVSEVEMLSQERIKVQGSPLKLFSSG
LKLKSGKQKGGGGDEPEGEYHQIQTSPESADEQKGESSAPSEPEEIALCLE
KQPSAPQAEAEAGTSDGKREGITPWAFKQVTPKRVPRPSDEKDEILDVY
KSATLSSTAGQDVEAVGEEORSEPRVDTVSWEALICVSGSKRRARKAS
SSDDEGPTLKGDAEASQKADALPASTODQOAHGSSSPAGSPSEGEV
STWSEFRLVTPRKSKLERRAEDSGAEQLASEIPEPSRESWSIKKFIPIGRKKR
ADGQEOAEVDEGPEINEDDPVAVPVPVSEYDAVEREKLEAQAQENVLPLQKG
AYVSELSKTLVHTVSAVIDTTRAVTSAEERSPSWISASMTPELHAEGVATPPVG
EYTKEDITAEATPALAQLPGGKAHDHIVTSEVDFSTSEAVTAETTEALRAEELTEA
SGAEETDMVSAVLSQSDSPDTEATPVQEVGMLDTEQORTQVLOAVADKVK
EDSQVPATQLORAGPKALEKEVEEEDSEVLATEKEDVPEVPQVQAEETRLAQS
ETVOATPESLEVPEDVDVDRATTCQVIKHQOLMEQAVAPESSELTLDSEITNGPLA
DSDTNGTQDDETVDSDNSNAIAVQSQVTEBEERAAATGEPSPSPQAEERERK
PGRVLEPTQALAAQAVPILAKAEGVQGEAGQFDGKVKDQGCCKVLEVPVHTGPS
OKTADLTRDSEVMEVARCOETSENEQSI SPEKREMGTDVKEETETKTEQASEHEQ
ETAAPHEGTHPKPVLITADPHSERGKALGSLGSPSLPDQKADCIQVQVSSDTPV
TOTTBVAKVEETVATSEMDSELECAQSLPAELSETGGYGTQLHGDEDTVPOGES
QAESPIIVTPAPESILHSDLOREVSASQKQSDENKDPDAGDRAKESAREKILR
APPEILESKNKIVQSVIQTAVQDFARTETAPETHASDLQNVPMQADSQGAQM
LDKDSLDQVSPQDGLTQVAGLAVSDSEMSKASEMITTLAVESAVSEVSKL
PLQCKDEKHAADGPHQSLAKAEADASGLTKESPDTPNGPLTBEQDALKEMNKAQ
TEEDDLOEPKGLTSES"
BASE COUNT 1840 a 1447 c 1779 g 1131 t
ORIGIN

Query Match 74.0%; Score 4559.6; DB 10; Length 6197;
Best Local Similarity 86.6%; Pred. No. 0;
Matches 5375; Conservative 0; Mismatches 644; Indels 189; Gaps 24;

QY 4 GCCGGCGGAGTAGAGCCACTGAGCCATGGCGGAGGAGGAGTCCACCGAGCAGGGA 63
DB 111 GCCGGCGGAGTAGAGGAGCCACTGAGCCATGGGTGCGAGGAGTCCACCGAGCAGGGA 170
QY 64 GCCCGGAGCAGCGCGGGGAGCGACACCGCGGAGCGAGCTGGTCTCAGTGGCCATGGGC 123
DB 171 GCCCGGAGCAGCGCGGGGAGCGACACCGCGGAGCGAGCTGGAGCTCAGTGGCCATGGGC 230

QY 124 CCGCAGCTGAAGCCTCGGAGCAGCTGGAGACCCCGCCGACGCGGACCCCGCCACCAAGC 183
DB 231 CCGCAGCGGAAGCGTGGGAGCAGCTGGAGATCCCGCTGACGCGGACCCCGCCACCAAGC 230
QY 184 TCCCACAGAAAGATGCGCAGCTGTCTTCTGTCAACGGCTAGCTGAACAAAGAGATGTCC 243
DB 291 TCCCACAGAAAGATGCTAGCTGTCTGCGCTCAATGGTGTAGCTGAACAAAGAGATGTCC 350
QY 244 ATGTCGAAGAGGAAAAACGAGGAGGCGAGGAGGAAAGTCTGTTGATGAGGATGTTGGAC 303
DB 351 ACGTCCAAGAGGAAAGCCAGGATGGCAAGAGGAAGTCACTGTTGAAGATGTTGGAC 410
QY 304 AGCGAGAGTCAGAGATGTGAGAGAAAAAGACCGAGTTGAAGAAATGGCGGCCAATCCA 363
DB 411 AGAGAGAGTCAGAGATGTGAAAGAAAAAGACCGAGCTAAAGAAATGGCAGCCAGTTCCA 470
QY 364 CAGCTGTGAAGATATACAAAGGATGGCAGGAGGAGACATCAGAAAAATTAATTGAACAGA 423
DB 471 CAGTTGTTGAAGATATACAAAGGACGAGGAGGAGAAACACCGGAAATTAATCGAACAGA 530
QY 424 TCCCTGCTTTCAGAAAAAATGTGGAAGAAATGGTACAGCTTGTGAGTCCAGGCTAATG 483
DB 531 TCCCTGCTTTCAGAGACCAATGTGGAAGAAATGGCGCAGGCTGTGAGTCCCAAGCTAATG 590
QY 484 ATGTTGGCTTCAAGAAAGTATTAAATTTTGTGGTTTTTAAATTCACGGTGAAGAGATA 543
DB 591 ACCTGGCTTTCAGAAAGGATTAAATTTTGTGGTTTTTAAATTCACGGTGAAGAGATA 650
QY 544 AAAATCAAAAGTCAGATCTGTCCAACTACTCCTCAAGGAGCCAGTGTGGAGACTGCCGTGGAG 603
DB 651 AAACGAAAGTCAGATACCGTCCAGTACTCCTGTCAAGAAAGGATGAAGCGCAAGGGG 710
QY 604 CAGAAACCTCTGTGCGAGCTGGAGACCAACGAGGAGCCAGTGTGGAGACTGCCGTGGAG 663
DB 711 CAGAAACCTCCGTGCGAGCAGGAGACCAACGAGAGCCCGAGTGGAGA---CCGTGGCG 767
QY 664 AGTCAGCTCCAAAGAAAGTGTAGCTGAAGCAATCCACAGAAAGCAAGAGCACCCTGA 723
DB 768 AATCAGCTCCAAAGAAAGTGTAGCTGAAGCAATCCACAGAAAGCAAGAGCACCCTGA 827
QY 724 AGCAAGAACAGAGCAGCAGCAAGAAATCCCCCTTCAAGCCCAATCTGATCAAGCGCTGAGG 783
DB 828 AGCAAGAACAGAGCAGCAGCAAGAAATTCCTTCAAGCCCAATCTGATCAAGCGCAGG 887
QY 784 AAGAACCCAAAGATGAAGGAGAGAGAAAAACAAGAGAAAGAGCCCAACCAAGTCCCCAGAAT 843
DB 888 AAGAACCCAAAGATGAAGGAGAGAGAAAAACAAGAGAAAGAGAAACCACTCAAGCCCTAGAT 947
QY 844 CCCCAGAGAGCCAGCTCAACAGTGAAGAACATCTTCTTCAAGAAAGTCTTCACTCAGC 903
DB 948 CTCGACACGAGCCCTGTGAGCAATGAGACCAACATCTTCTTCAAGAAATCTTCACTCAGC 1007
QY 904 GTTGGCGCGCTGGCGCAGAGACCAAGCTTCAAGAAATCAAAAGAGGATGATCTGGAAA 963
DB 1008 GGTGGCGCGCTGGCGCAGAGACCAAGCTTCAAGAAATCAAAAGAGGATGATCTGGAAA 1067
QY 964 CTGCGCAG 1023
DB 1068 CTTCCGAG 1127
QY 1024 CAGAGCCAGCC---TCGGAGGAGCAGGAGCCCGGAGAGACACAGACAGGCGCAGGTTGT 1080
DB 1128 CAGAGCCAGCCCGCAGCGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1187
QY 1081 CAGCAGACTACAGAGAGTGGAGCTGCCCTTTTGAAGACCAAGTTGTGTGACCTGGAGGAT 1140
DB 1188 CAGCGGACTATCAGAGAGTGGAGTTGCCCTTTTGAAGACCAAGTTCGCTGACCTGGAGGAT 1247
QY 1141 CGTCAGAGAGAGAGTGTCTCTCTTTGGCAACGGAAGTGTGTTGATGAGAGATGGAAGCCC 1200
DB 1248 TGTCT---GGAGAGTGTGCTCTCTTTGGCAACGGAAGTGTGTTGATGAGAGATGGAAGCCC 1304
QY 1201 ACCAAGAGTGTGTGAGAGGTCACGCTGAGCACCGTGGAGAGACAGAGGAGGAGCAGG 1260

Db	1305	ACCAAGAGTTGTTCCAGAGGTCACAGTGAAGCCCTGGAGAAAGATGACGAAAGGCA--	1362
Oy	1261	GAGGAGGAGGAGGCTGAAGGGGGCGTGGGTAGAAAGAAACAGAGAAATCTTGGCCCC	1320
Db	1363	-AGGAGACACAGGTGGGAAGGGGATGTGTGTGTGGAAGATCGGAGAAATCTTGGCCCC	1421
Oy	1321	CTGAGAAACTGAGCTGAGCCCCCAGAGAGTCCCCCAGGAAGCTGAGCTGTGAGAGCTGA	1380
Db	1422	CTGAGAAACTGGCTGAGACCCAGAGAGTCCCCCAGGAAGCTGAGCTGTGAGAGAGCTGA	1481
Oy	1381	TGAAGAGCAGAGATGTGTGTCTCTGAGAGAGAACCAACTCAACTGACAGACCTTAAGTC	1440
Db	1482	TGAAAGCCAAAGAAAGTACGTCCTCTGGGGGTGACATCACTCAGCTGACAGAACTTAAGTC	1541
Oy	1441	CTGAAGAGAAAGCGCTGGCCCAACCCAGAAAGCAATTGCACTGAGGTGAGATCTGTG	1500
Db	1542	CTGAAGAGAAAGTCTTACCCAAACCCCGAAGGCAATTGTCACTGAGGTGAGATCTGTG	1601
Oy	1501	CTCTCAGAGAAAGATCAAGGTACAGGGAAGTCCCTTGAAGAACTCTTCAGTAGCTCAG	1560
Db	1602	CTCTCAGAGAGAAATCAAGGTACAGGGAAGTCCCTTGAAGAACTCTTCAGTAGCTCAG	1661
Oy	1561	GCTTAAAGAACTGTCTGGAGAGACAGAGAGGGA--ACAGAGAGGTGGGGAGAGC	1617
Db	1662	GCTTAAAGAACTCTCGGGAGAGACAGAGAGGGAAGAGAGAGAGAGGGGGAGATG	1721
Oy	1618	AAGAGCTGAGAAATCAACAATTCACACCGAATCCCGAGAGAGTGTGATGAGACA	1677
Db	1722	AAGAGCTGAGAAATCAACAATTCACACCGAATCCCGAGAGAGTGTGATGAGACAGA	1781
Oy	1678	AAGGAGAGAGCTTGGCGTCCGCCGAGAGGCTTGAAGAGACAAGTGTGAGAAAG	1737
Db	1782	AAGGAGAGAGCTTGGCGTCCGCCGAGAGGCTTGAAGAGAGTGGCGTGTGAGAGAG	1841
Oy	1738	GCGCGCTGAAAGCACTCCAGAGATGGGGAGCTGAGAGAGACTACTTCCGATGAGAGAG	1797
Db	1842	GCGCATGCGAAGCACTCCAGAGAGCGGAAGCTGAGAGAGAGCACTTCCAGAGGAGAGA	1901
Oy	1798	AGAAGAGAAAGGATCACTCCCGGGCACTCTTCAAAAAGATGTGTACACCCAAAGAAC	1857
Db	1902	AGAAAGAGAAAGGATCACTCCCGGGCACTCTTCAAAAAGATGTGTACACCCAAAGAAC	1961
Oy	1858	GAGTCCGAAGACCTTCTGAGAGTGAACAAGAGAGAGAGCTGTGAGAAAGTCAAGAGCGCA	1917
Db	1962	GAGTCCGAAGACCTTCTGAGAGAGAGCAAGAGAGAGAGAGAGTGAAGTCAAGAGTCCCA	2021
Oy	1918	CCTTGTCTTCACTGATAGCAAGTGTCAAGAAATGCAAGTGAAGTCAAAACTGTGTGTG	1977
Db	2022	CCTTGTCTTCACTGAGAGAGCAAGGCGTCTGGAATGCGAGATGAGTCAAGAGCGTGTGCG	2081
Oy	1978	AGGAACAAAGCCAGAGAGAACAAAGCGTGAAGGTGAGTACTCTCAAGTGTCTTGGGAACAC	2037
Db	2082	AGGAGCAAAAGTGAAGAGAGCCAAAGCGCAGGGTGAATCTTCAAGTGTCTTGGAGCGCT	2141
Oy	2038	TGATTTGTGTCCGATATCAAGAGAGAGCAAGAGAGGATCTCTTCAGATGATGAAG	2097
Db	2142	TGATTTGTGTCCGATCTTCAAGAGAGAGCGAGAGAGGATCTCTTCAGATGATGAAG	2201
Oy	2098	GAGGCGCCAGAGACTGGAGAGGGGACGTCAAGAGCAAGAGAGCGCAAGCAAGCAAAAG	2157
Db	2202	GAGGCGCCAGAGACTGGAGAGGGGATGGCCACAGAGCGGAGAGGCTTGAAGCAAAAG	2261
Oy	2158	AAGCGCGAACAAGAGCTGTCTCTGCGAGACCCAGAGAGCGAGACCAAGCGCAAGAAAGTT	2217
Db	2262	AAG-----CAGATGTCTTCTGCGAGACCCAGAGAAACAAGACCAAGCGCAAGAAAGTT	2315
Oy	2218	CCTCAACCGAGCCAGCGGAGAGCCCTTCCGAGAGGGAGAGTGTCTCCACTTGGAGTCAAT	2277
Db	2316	CCTCAACCGAGCCAGCTGGAAGCCCTTCTGAAGGGAGGGGTCTTCACTCGGAGTCAAT	2375
Oy	2278	TTTAAAGATTAGTCACTTCAAGAAAAAATCAAGTCAAAACTGGAAGAGAACCCGAG	2337

Db	2376	TTTAGAGATTGTTGACCTCCAGAAAAAAATCCAAAGTCAAAACTG99AAGAGAGCCGAG	2435
Qy	2338	ACTCTAGTGAAGCAGTTGTTCTCACTGATCTGAACCGAGTGAAGAAATCTTGGGTTT	2397
Db	2436	ACTCCGGTGCAGAGCAGTTTGCCCTCCGAGATCGAACCAAGTAGAGAGAAATCTTGGGTTT	2495
Qy	2338	CCATTAAAGAAATTCATCCCGGGAAGGAGGGAAGAAAGGCGCAGCGGGAAGCAAG	2457
Db	2496	CCATTAAAGAAATTTATCTTGAGCGGCGGAAGAAAGGCGCAGATGGAGACAGAACAG	2555
Qy	2458	CCACTGTGAAAGACTCAGAGGCGCAGTGAATAAATGAAGACGACTATATGTCCAGCCG	2517
Db	2556	CCGCGGTTGAAGACTCCGGGGCCAGAGAGATCAATAGAGACGACCCCGACGTCCAGCTG	2615
Qy	2518	TCGTGCTCTGTCTGAGTATATATGCACTGAGAGGAGAGATGGA-----AGCCC	2568
Db	2616	TGTGCTCTGTGTTGATGATGATGCGGTAGAGAGAGAAACCTGGAACCGACGAGCTTC	2675
Qy	2569	AGGGGAATACGAGAGCTGCCCCAGCTGCGGGGCGTGTATAGTGTCCGAGAGCTCACTA	2628
Db	2676	AGGAAGACGTGAGGTGCCCCAGCTGAAGGGGGCGTGTATATGTCTGAGAGCTTAGTA	2735
Qy	2629	AGACTCTGGTCCACTGTGAGTGTGCGAGTCAATTGATGGGACCCAGGCGACTGCACATG	2688
Db	2736	AACACTGTGTTCAACTGTGAGTGTGCGGTCATTGATGGGACCAAGGCGACTGCACATG	2795
Qy	2689	TCGAAAGCGGATCTCTCTGTGGATATCCGCTTCGTTAACAGACTCTTTGACACACAG	2748
Db	2796	CCGAAGAGCGGTCCTCTGTGAGATATCTGCTTCATGACAGAACTCTTGGAGCCGAG	2855
Qy	2749	CGGAGAAAGCCATGCGACCTGTGTAAGAGTCACTGAAAAAGACATCATTCGAGAAAGAA	2808
Db	2856	AGGGAGTGGCCACACCGCTGTGGAAGAGTCACTGAAAAAGACATCATGAGAAAGCA	2915
Qy	2809	CTCCGTGTCTCAACCAGACGTTACCGAGAGGGTAAAGATGCCATACGACATGTCTACCA	2868
Db	2916	CTCTGTGCACTCCGCCAGACTTTACAGGGGGGAAAGATGCCATACGACATAGTCAACA	2975
Qy	2869	GTGAATGTGATTTCACTTCAGAGCTGTGACGCCACAGACACTTCAGAGCTCTCCGTA	2928
Db	2976	GTGAATGTGATTTTACTCTCAGAAAGCATGACAGCGCAGAAACCAAGAGGCGCTCCGG	3035
Qy	2929	CTGAAGAACTTACCGAAGATGCGGGGCCGAAGACACACATATGTGTCCGAGTTT	2988
Db	3036	CTGAAGAACTTACCGAAGATCAGGGCGGAAGAGACACACATATGTGTCCGAGTTT	3095
Qy	2989	CCGAGTGACTGACTCCCGACAGACCAACAGAGAAAGCCACCAGTTCCAGAGTAGAGA	3048
Db	3096	CCGAGTGTTCGACTCCCGGACACCAACAGAGAAAGCCACCAGTTCCAGAGTAGAGG	3155
Qy	3049	GTGTGTGTGTAATACAGAAAGAGAGAGCGCCAGACCGAGGCCATCTCCAAGCCGTTG	3108
Db	3156	GTGTGTGTGTAATACAGAAAGAGAGAGCGCCAGACCGAGGCCGCTCCAAGCCGTTG	3215
Qy	3109	CAGACAAAGTGAAGAGAGTCCGAGGTCCCTGCAACCCAGCTGTGCGAAGAACCGGGT	3168
Db	3216	CAGACAAAGTGAAGAGAGTCCGAGGTCCCTGCAACCCAGCTGTGCGAAGAGCGGCC	3275
Qy	3169	CAAAACCACTGGAAGAGTGTGAAGAGTGAAGAGACTCCGAAAGTGTGGCTTCCGAGA	3228
Db	3276	CGAAAGCACTGGAAGAGTGTGAAGAGTGAAGAGACTCCGAGGTGTGGCTTCCGAGA	3335
Qy	3229	AAGAGAGAGCGTTATGCGGAAGAGACCCGTCAGAGAACTGGAAGCTGAGACTTTGAC	3288
Db	3336	AAGAGAGAGTGTGTGTGCGGAAGAGACCCGTGAGGAAGCTGAAACTGAGCAATCTTGAC	3395
Qy	3289	AGGGCTCTGAGACTGGAAGAGCTACTCTCAGAGAGCTTTGAAGTTCTTGAAGTCAAGGACG	3348
Db	3396	AGGGCTCTCAGACTGTACAGGCTTACCCAGAGAGCTTTGAAGTTCTTGAAGTCAAGAGG	3455
Qy	3349	ATGTAGACCATGTGCGCAGGTGCAAGTTATCAAGTCCAGACAGCTGATGGAACAGAGCGG	3408
Db	3456	ATGTAGACCGTGCACACATGCGAAGTTATCAAGCACAGAGCTGATGGAACAGAGCTG	3515


```

Db      5561  GTATATTTTCTTCTAGGTTTACGACATGCTTTTGTATTTATGACATATATATGSGT 5620
Qy      5593  GTGAGCCATGACGACGCTTTGAAAAGCTCCAGAGCTCACTGTAACCTGACGAAACAG 5652
Db      5621  GTGAGCCATGACGACGCTTTGAGAAGCCCAAGCTCACTGTAACCTGACGAAACAG 5680
Qy      5653  ----ATTAATCTCTGCGAAGAGACAGCTTTTAAAGTTTACTGATGCTTAAT 5708
Db      5681  AGAATATACATCTCGGACGAGAGATACAGTC-TTTTTAAAGTTTACTGATGCTTAAC 5739
Qy      5709  CTGCGGCTCTAGTCCCTGTAAGAGGTTGTTTCTATGACAGCAGCAGTCAAGATA 5768
Db      5740  CTGCGGCTCTAGTCCCTGTAAGAGGTTGTTTCTATGACAGCAGTCAAGATA 5799
Qy      5769  AAAACCCATTTTGAAGATCCAGAGTGTCCCAATATTAACATGATTTTCCCTTT 5828
Db      5800  AAAATCTCATTTTGAAGATCCAGAGTGTCCCAATATTAACATGATTTTCTTTT 5859
Qy      5829  TTGCTAATCCAGTCCAGTGGAAAAGAGTCTCTCTGTGTGATTAAGCCCTGCTCT 5888
Db      5860  TTTTATATCCAGTCCAGTGGAAAAGAGTCTCTTATGATGATTAAGCCCTGCTCT 5919
Qy      5889  TATATATATGACCAATAGTGTGCTTAAGGCGCATGAGATGTTTCTATGACAGGAA 5948
Db      5920  TAAACATATGACCAATAGTGTGCTTAAGGCGCATGAGATGTTTCTATGACAGGAA 5978
Qy      5949  TCTGTATGACGTTTGTGATGTTGATGTTGATGATGATGATGATGATGATGATG 6008
Db      5979  TTTGTTGTTACTTTT-GGATTGTACTCTTCTAGGCTGACCGAATTCATATGACGAT 6037
Qy      6009  AAGTAGTCTCTCTCTTCAAGATGATTTTGTATGATGATGATGATGATGATGATG 6068
Db      6038  AAGTAGACCTCTCTCTTCAAGATGATTTTGTATGATGATGATGATGATGATGATG 6097
Qy      6069  ATCTGTGCCCCCTTCTTAAAGCAATGTTGATTAATGTTCTTGGATTAATGATGAT 6127
Db      6098  ATCTGTGCCCCCTTCTTAAAGCAATGTTGATTAATGTTCTTGGATTAATGATGAT 6157
Qy      6128  TGACAACGATTTAAATAACATATTTG 6155
Db      6158  TGACAACGATTTCAATAAATATTTG 6185

```

```

RESULT 5
AB020886 6195 bp mRNA linear ROD 14-APR-1999
LOCUS AB020886 Mus musculus mRNA for Ssecks, complete cds.
DEFINITION AB020886
ACCESSION AB020886.1 GI:4589831
VERSION AB020886.1
KEYWORDS Ssecks.
SOURCE Mus musculus (strain: BALB/c, isolate: 2001) 9-12 weeks male testis
tissue, lib: Clontech mouse testis library cDNA to mRNA, clone: 133.5.
ORGANISM Mus musculus
REFERENCE 1 (sites)
AUTHORS Okita, K., Kitamura, H., Iwanaga, T., Morimatsu, M. and Saito, M.
TITLE Isolation and characterization of mouse Ssecks
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 6195)
REFERENCE Okita, K., Kitamura, H., Saito, M. and Morimatsu, M.
AUTHORS Direct Submission
TITLE Submitted (08-DEC-1998) Keisuke Okita, School of Veterinary
JOURNAL Medicine, Hokkaido University, Laboratory of Biochemistry, N18, W9,
Kite-ku, Sapporo, Hokkaido 060-0818, Japan
(E-mail: v08008@vet.ec.hokudai.ac.jp, Tel: 81-11-706-5206,
Fax: 81-11-757-0703)
COMMENT Sequence updated (10-Mar-1999).
FEATURES
location/Qualifiers
1..6195
/organism="Mus musculus"
/strain="BALB/c"

```

```

REFERENCE 1 (sites)
AUTHORS Okita, K., Kitamura, H., Iwanaga, T., Morimatsu, M. and Saito, M.
TITLE Isolation and characterization of mouse Ssecks
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 6195)
REFERENCE Okita, K., Kitamura, H., Saito, M. and Morimatsu, M.
AUTHORS Direct Submission
TITLE Submitted (08-DEC-1998) Keisuke Okita, School of Veterinary
JOURNAL Medicine, Hokkaido University, Laboratory of Biochemistry, N18, W9,
Kite-ku, Sapporo, Hokkaido 060-0818, Japan
(E-mail: v08008@vet.ec.hokudai.ac.jp, Tel: 81-11-706-5206,
Fax: 81-11-757-0703)
COMMENT Sequence updated (10-Mar-1999).
FEATURES
location/Qualifiers
1..6195
/organism="Mus musculus"
/strain="BALB/c"

```

```

/isolate="200"
/db_xref="taxon:10090"
/clone="133.5"
/sex="male"
/tissue_type="testis"
/dev_stage="9-12 weeks"
/tissue_lib="Clontech mouse testis library"
1..6195
/gene="ssecks"
140..5194
/gene="ssecks"
/note="Sic Suppressed C Kinase Substrate"
/product="Ssecks"
/protein_id="BA076894.1"
/db_xref="GI:4589832"
/translation="MGAGSTGEQSPPEQAPSEDTPELEISGHGPAABAGAADPAD
ADPATLTPKNGQLSANGVAREQEDVHVEESQDQGEETVDEVOQRESEDEKEDR
AKEMASSTVEDITKDEQETPEIIEQIPASESVEEMAOAESQANDVGFKVFVF
VGFKFTYKKDKNEKSDPTVOLITVKDEGEGEASVAGDHQDEGVETVGSASSEL
KQSTEROEGLKQANOSTEPILOAESGQGEEREADGEMREKREPKPLESPSPVS
NETTSTFKFTFGMAGMKRTISFKPKEDDLTESEKKEQAEKDEDEGCTETPA
ABOEPAEGTDARLSADYEKVELPLEDQVLEALSEKCAPLATEVDEKTAHOBV
VAEVHVSIVEMKMTGQGAEEGVVVEGSESLPEKLAETEVDEQAEVPELMKT
KEVCSGDHTOLTDLSPBEKMLPKPEGVISVEMLSQERIKVQSPPLKPLPSBSG
LKLSGKKQKRGKGGGDEPEYOHIOETESPESADDEKESASSPEPERIACLE
KGPSEAOEABEABEGATSDSEKREGITTPASRKMYTPKKRVRRPESKRELDV
KSATLSTESIASGEMODEVAVABEBOSEBPKRRVDTSVEMALICYSKKARAS
SSDEGEPRLTGDDHRAEASAKDEKDALPASTOEDQHGSSSPBEPSEBGEV
STWESFKRLTPPKRSKSLERAEBSQALASIEPSEBWSIKKPIPKRRKR
ADGKOEAIVEDSGPEINEDDPVAVVPLEDAVERKLEAQRQENVEVLPOLG
AVVSEELSTIVHTVSVAVIDGTRATSAEERSPSMISAMTEPLEHAGVAVLPFG
EYERKDTTATATPALAOTLPQGDADHDITYSVDTPSEVTAETTAETTAETTA
SGAETTDWVSAYSQSDSPDTEATPVEVEKGMLEDEQEROTAVLOVADTK
EDSQVPATQTLORAGPALEKVEVEEDSVLATEKEDVAPESPVEATEHIAQDS
EYQVATPESLEVEPTEVDVRAATCOVITKHQOIMEQAVABESSSTLDSBTNGSTPLA
DSPTPNTODEEDSDSNALIAVAKOSQVTEBAAAQTEGSTPSPFAOEHRK
GRDVELEPQOLAAGAPILAKAEVGGEGAGGDEKVDGOCVKELEVPVHTGPS
OKTRDLTRDSEWENACQCTESNEBOSIPREKREKTDVKEKTEKTAOASEHNO
ETAPVEHGHHPKVLTAIDMPSHREGALSLSESPBLPQDRAKICIEVOYSSDTTV
TQTEAVKVEEIVATSEMDSELESCAQSILPAKUSSETGYGLQGEDVTPGPGS
QAESIPITVTPAPESILHSDLOREVSASQKORSBDEKPPAGPDAKESAAARKILR
AEPILEESKNKIQSVIOTAVDQPARFETAPETASLDLPVMAQSCQAQM
LDKESDLOVSPODGTLSAVAOEGIAVSDESEKASEMITTLAVASVSEVEKL
PLOCDEKHEAAGROHOSLAKAEADASGNLTRESPTNPKLTEBDALKENKNO
TEBDLQEPKDLTES"
16195
/gene="ssecks"
/note="20 A nucleotides"
BASE COUNT 1838 a 1451 c 1777 g 1129 t
ORIGIN
Query Match 74.0%; Score 458; DB 10; Length 6195;
Best Local Similarity 86.6%; Pred. No. 0;
Matches 5374; Conservative 0; Mismatches 645; Indels 189; Gaps 24;

```

```

Qy      4  GCCGCGGAGTAGAAGAGCCACTAGCCATGCGCGGACGATTCACCGACGCGGCA 63
Db      111 GCCGCGAGAGTAGAAGAGCCACTAGCCATGCGCGGACGATTCACCGACGCGGCA 170
Qy      64  GCCCGAGAGCGCGCGGAGCGACACGCGAGGAGCTGTGCTAGTGCGCATGGGC 123
Db      171  GCCCGAGAGCGCGCGGAGCGACACGCGAGGAGCTGTGCTAGTGCGCATGGGC 230
Qy      124  CCGCAGCTGAAGCTCTGCGAGCAGCTGAGACCCCGCAGCGACCGACCCCGCACCAAGC 183
Db      231  CCGCAGCGGAACGCTGCGAGCAGCTGAGACCCCGCAGCGACCCCGCACCAAGC 290
Qy      184  TCCACAGAAAGATGCGACGCTGTCTTCTGTAAAGCGGTAGCTGAACAAGAGATGTC 243
Db      291  TCCACAGAAAGATGCTGCTGTCTGTCTGTAAAGCGGTAGCTGAACAAGAGATGTC 350
Qy      244  ATGTCCAGAGGAAACCGACGAGGCGCAGGAGAAAGATGCTGTATGAGATGTTGAC 303

```

[illegible]

Db	1422	CTGAAACCTGGCTGAGACCCAGGAGGTCCTCCCGAGGAAGCTGAGCCTGTGTGGAGGAGCTGA	148
Qy	1381	TGAAGAGCAGAGAGATGTGTCTCTGGAGGAGACCAACACTCAACTGACAGACCTAAAGTC	1440
Db	1482	TGAAGACCAAGAAGTATGCGTCTCTGGGGGTGACCATACTCAGCTGACAGATCTAAGTC	1541
Qy	1441	CTGAAGAGAGACGCTGCCCAACACCCAGAAAGGCATTGTTCAGTAGAGGTGGAGATGCTGT	1500
Db	1542	CTGAAGAGAGATGCTACCCAAACACCCCGAAGGCATTGTTCAGTCAGGTGGAGATGCTGT	1601
Qy	1501	CCTCTCAGGAAGAATCAAGGTACAGGGAAGTCCCTTCGAGAAATCTTTCAGTAGCTCAG	1560
Db	1602	CCTCTCAGSAGAAATCAAGGTACAGGGAAGTCCCTTCGAGAAAGCTCTTTCAGCAGTTCGG	1661
Qy	1561	GCTTAAAGAAAGCTGTCTGGGAAGACGACGAGGAA---ACGAGGAGGTGGGGAGACG	1617
Db	1662	GCTTAAAGAGCTCTCCGGAGAACGACAGAGGAGAGAGGAGAGGCGGGGAGATG	1721
Qy	1618	AAGAGCTTGAGAAATCCAAACATTCACCCGAATTCGCCAGAGAGTCTGATGAGCAGA	1677
Db	1722	AAGAGCCAGGAGAAATCCAAACATTCAAACCGAGTCCCGACAGAGTCTGCACGAGCAGA	1781
Qy	1678	AGGAGAGAGCTCTGGTGTCTCCCGAGGAGCCTGAGGAGACACGCTGCTCGAGGAAG	1737
Db	1782	AGGAGAGAGCTCTGCTCTCTCCCTTGAAGAGCCCGAGAGATCGCGTGTCTGGAGAAG	1841
Qy	1738	GGCCGCTTGAAGACCCACAGGATGGGGAAGCTGAGGAAGGAATCTTCCGATCGAGAGA	1797
Db	1842	GGCCATCGGAAGCACCCAGGAAGCGGAAGCTTGAGGAAGGACGACTTCCAGCGGAGAGA	1901
Qy	1798	AGAAGAGAAAGGATCACTCTCTGGGCATCTTCAAAAAGATGGTGAACCCCAAGAAAC	1857
Db	1902	AGAAAAGGAAGGATCACCCCTGGGCATCTTCAAAAAGATGGTGAACCCCAAGAAAC	1961
Qy	1858	GGGTCCGAAGACCTTCTGAGATGACAAAGGAGAGACCTGGAGNAGGTCAAGAGCGCCA	1917
Db	1962	GGGTCCGAAGACCTTCTGAGAGCGACAAGGAAGAGAGCTGGATNAGGTCAAGAGTGC	2021
Qy	1918	CTTTGTCTCTCACTGATAGCACAGTGTCAAGAAATCAAGATGMAGTCAAACTGTTGGT	1977
Db	2022	CTTTGTCTCTCAGGAGACACGGCGTCTTGGAAATGAGATGAGGTGAGACGGTGGCG	2081
Qy	1978	AGGAACAAAACCCAGAGGAACCAAGCGTGGGTGGATCTTCACTGTCTTTGGGAAGCAC	2037
Db	2082	AGGAGCAAGGTGAGAGGAGCCAAAGCGCAGGCTGGATACTTCACTGTCTTTGGGAGCGT	2141
Qy	2038	TGATTTGTGTCGGATCATCAAGAGAGAGCAAGGAAGGCATCTCTTTCAGATGATGAAG	2097
Db	2142	TGATTTGTGTCGGATCGTCCAAAGAGAGACGAGGAAGGCATCTCTTTCAGATGATGAAG	2201
Qy	2098	GAGGGCCAGGACACTGGGAGGGGACAGTCAAGAGCAGAGGAGCGCCAGCAAGACAAG	2157
Db	2202	GAGGCCAAGAACACTGGGAGGGGATGGCCACAGAGCGGAGAGGCTAGCAAAACAAG	2261
Qy	2158	AAGCCGGAACAGACGCTGTTCTCTGCAGCACCCAGGAGCAGGACCAAGCGCAAGGAAGTT	2217
Db	2262	AAG-----CAGATGCTCTCTCTGCAGCACCCAGGAACCAAGCAAGCGCACGGAAGTT	2315
Qy	2218	CCTCACCGGAGCAGCGGGAAGCCCTTCGAAAGGGGAAGGTGTCTCTCACTTTGGGAGTCAT	2277
Db	2316	CCTCACCGAGCAGCTGGGAAGCCCTTCTTGAAGGGGAGGGGCTCTTCCACCTTGGGAGTCAT	2375
Qy	2278	TTAAAGATTAGTCACTCCAGAAAAAATCAAGTCAAACTCGAGAGAAAGCCGAG	2337
Db	2376	TTAAGATTAGTCACTCCAGAAAAAATCAAGTCAAACTCGAGAGAGAGCCGAG	2435
Qy	2338	ACTCTAGTGTAGAGCAGTTGTCTCACTTGAGATCGAATCGAATAGAGAAGAAATCTTTGGGTTT	2397
Db	2436	ACTCGGTGCAGAGCAGTTGGCCTCCGAGATCGAACCAAGTAGAGAGGAATCTTTGGGTTT	2495
Qy	2398	CCATTAAAGAAATTCATCCCGGACGGCGGAAGAAAGGGCAGACGGGAAGCAAGAACAG	2457
Db	2496	CCATTAAAGAAATTTATCTCTGACGGCGGAAGAAAGGGCAGATGGGAAGCAAGACAGG	2555


```

Massey, E., Mawhinney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M.,
Miner, G., Miner, Z., Mitchell, T., Monabadi, K., Morgan, M., Morris, S.,
Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
Nguyen, N., Nickerson, E., Nwokweto, S., Ogun, M., Okunolu, G.,
Ogunyemi, N., Oviedo, R., Pace, A., Payton, B., Peery, M., Ren, Y.,
Peters, L., Pickens, R., Primus, E., Pu, L., Quiles, M., Ren, Y.,
Rivas, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G.,
Scherer, S., Scott, G., Shen, H., Shooshkat, N., Sisson, I.,
Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Swatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tang, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Umani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Warlington, S.,
Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G. and Gibbs, R.
Direct Submission
Unpublished
2 (bases 1 to 178954)
Worley, K.C.
Direct Submission
Submitted (22-MAR-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 178954)
Worley, K.C.
Direct Submission
Submitted (13-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 12, 2002 this sequence version replaced gi:21239989.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu

----- Project Information
Center project name: GPHP
Center clone name: CH230-171G16
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap, version 0.990329
Consensus quality: 142023 bases at least Q40
Consensus quality: 146696 bases at least Q30
Consensus quality: 150215 bases at least Q20

-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 53 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1459: contig of 1459 bp in length
* 1460 1559: gap of unknown length
* 1560 2931: contig of 1372 bp in length
* 2932 3031: gap of unknown length
* 3032 4420: contig of 1389 bp in length
* 4421 4520: gap of unknown length
* 4521 5800: contig of 1280 bp in length
* 5801 5900: gap of unknown length
* 5901 7559: contig of 1659 bp in length
* 7560 7659: gap of unknown length
* 7660 8887: contig of 1228 bp in length
* 8888 8987: gap of unknown length
* 8988 10297: contig of 1310 bp in length
* 10298 10397: gap of unknown length
* 10398 12090: contig of 1693 bp in length
* 12091 12190: gap of unknown length

12191 13522: contig of 1332 bp in length
13523 13622: gap of unknown length
13623 15629: contig of 2007 bp in length
15630 15729: gap of unknown length
15730 17540: contig of 1811 bp in length
17541 17640: gap of unknown length
17641 19819: contig of 2179 bp in length
19820 19919: gap of unknown length
19920 21261: contig of 1342 bp in length
21262 21361: gap of unknown length
21362 23786: contig of 2424 bp in length
23786 23865: gap of unknown length
23866 25562: contig of 1677 bp in length
25563 25662: gap of unknown length
25663 27442: contig of 1780 bp in length
27443 27542: gap of unknown length
27543 30415: contig of 2873 bp in length
30416 30515: gap of unknown length
30516 31784: contig of 1269 bp in length
31785 31884: gap of unknown length
31885 34970: contig of 3086 bp in length
34971 35070: gap of unknown length
35071 36233: contig of 1163 bp in length
36234 36333: gap of unknown length
36334 39668: contig of 3335 bp in length
39669 39768: gap of unknown length
39769 42866: contig of 3098 bp in length
42867 42966: gap of unknown length
42967 46088: contig of 3122 bp in length
46089 46188: gap of unknown length
46189 49027: contig of 2839 bp in length
49028 49127: gap of unknown length
49128 51579: contig of 2452 bp in length
51580 51679: gap of unknown length
51680 55192: contig of 3513 bp in length
55193 55292: gap of unknown length
55293 58204: contig of 2912 bp in length
58205 58304: gap of unknown length
58305 60939: contig of 2635 bp in length
60940 61039: gap of unknown length
61040 64184: contig of 3145 bp in length
64185 64284: gap of unknown length
64285 67776: contig of 3492 bp in length
67777 67876: gap of unknown length
67877 72019: contig of 4143 bp in length
72020 72119: gap of unknown length
72120 75147: contig of 3028 bp in length
75148 75247: gap of unknown length
75248 78405: contig of 3158 bp in length
78406 78505: gap of unknown length
78506 82238: contig of 3733 bp in length
82239 82338: gap of unknown length
82339 85153: contig of 2815 bp in length
85154 85253: gap of unknown length
85254 87767: contig of 2514 bp in length
87768 87867: gap of unknown length
87868 90643: contig of 2776 bp in length
90644 90743: gap of unknown length
90744 94013: contig of 3270 bp in length
94014 94113: gap of unknown length
94114 98293: contig of 4180 bp in length
98294 98393: gap of unknown length
98394 101682: contig of 3289 bp in length
101683 101782: gap of unknown length
101783 104237: contig of 2455 bp in length
104238 104337: gap of unknown length
104338 108706: contig of 4369 bp in length
108707 108806: gap of unknown length
108807 113208: contig of 4402 bp in length
113209 117308: gap of unknown length
117309 117338: contig of 3930 bp in length
117339 122577: gap of unknown length
122578 122577: contig of 5239 bp in length

```

```
* 122578 122677: gap of unknown length
* 122678 122679: contig of 4102 bp in length
* 126780 126781: gap of unknown length
* 126880 126879: contig of 4928 bp in length
* 131808 131907: gap of unknown length
* 131908 138315: contig of 6408 bp in length
* 138316 138415: gap of unknown length
* 138416 143597: contig of 5182 bp in length
* 143598 143697: gap of unknown length
* 143698 152755: contig of 9058 bp in length
* 152756 152855: gap of unknown length
* 152856 157897: contig of 5042 bp in length
* 157898 157997: gap of unknown length
* 157998 167260: contig of 9263 bp in length
* 167261 167360: gap of unknown length
* 167361 178994: contig of 11594 bp in length.
FEATURES
  Location/Qualifiers
    Source          1..178954
      Query Match      72.1%; Score 4441.8; DB 2; Length 178954;
      Best Local Similarity 99.1%; Pred. No. 0;
      Matches 4517; Conservative 0; Mismatches 34; Indels 6; Gaps 5;
QY 297 GTTGGCAGCGAGAGTCAGAGATGTGAGAGAAAGACCCGAGTTGAAGAAATGGCGGCC 356
Db 174384 GTTGGCAGCGAGAGTCAGAGATGTGAGAGAAAGACCCGAGTTGAAGAAATGGCGGCC 174443
QY 357 AACTCCACAGCTGTTGAAGATATCACAAGAGTGGCAGGAGGAGACATCAGAAATAATT 416
Db 174444 AACTCCACAGCTGTTGAAGATATCACAAGAGTGGCAGGAGGAGACATCAGAAATAATT 174503
QY 417 GACAGATCCCTGCTCAGAAAACAATGTGAGAAATGTCACAGCTGCTGATGCCAG 476
Db 174504 GACAGATCCCTGCTCAGAAAACAATGTGAGAAATGTCACAGCTGCTGATGCCAG 174563
QY 477 GCTAATGATGTTGGCTTCAAGAAAGTATTTAAATTTCTGTTTTTAAATTCACGGTGAAG 536
Db 174564 GCTAATGATGTTGGCTTCAAGAAAGTATTTAAATTTCTGTTTTTAAATTCACGGTGAAG 174623
QY 537 AAGGATAAAATGAAAGTCAGATATCTGCTCAACTACTCTGTCAGAGAGGATGAAGGC 596
Db 174624 AAGGATAAAATGAAAGTCAGATATCTGCTCAACTACTCTGTCAGAGAGGATGAAGGC 174683
QY 597 GAAGGGCAGAGCCTCTGTGAGAGTGGAGACACACAGAGCCAGTGTGGAGACTGCC 656
Db 174684 GAAGGGCAGAGCCTCTGTGAGAGTGGAGACACACAGAGCCAGTGTGGAGACTGCC 174743
QY 657 GTCGAGAGTCAGATCCAAAGAAAGTGAGCTGAAGCAATCCACAGAGAACGAGAGGC 716
Db 174744 GTCGAGAGTCAGATCCAAAGAAAGTGAGCTGAAGCAATCCACAGAGAACGAGAGGC 174803
QY 717 ACCCTGAAGCAAGAACAGACAGACAGAGAAATCCCTTCAAGCCGAATCTGATCAAGCG 776
Db 174804 ACCCTGAAGCAAGAACAGACAGACAGAGAAATCCCTTCAAGCCGAATCTGATCAAGCG 174863
QY 777 GCTGAGAGAACCAAGATGAAGAGAGAGAAACACAGAGAAAGCCCAACAAGTCC 836
Db 174864 GCTGAGAGAACCAAGATGAAGAGAGAGAGAAACACAGAGAAAGCCCAACAAGTCC 174923
QY 837 CCAGAAATCCCGAGAGCCAGTCACAGTGAGACACATCTTCTTCAAGAGTTCTTC 896
Db 174924 CCAGAAATCCCGAGAGCCAGTCACAGTGAGACACATCTTCTTCAAGAGTTCTTC 174983
QY 897 ACTCACGGTTGGCGCGCTGGCGCAAGAACACAGCTTCAAGAAATCAAAAGAGGATGAT 956
Db 174984 ACTCACGGTTGGCGCGCTGGCGCAAGAACACAGCTTCAAGAAATCAAAAGAGGATGAT 175043
QY 957 CTGGAATCTCCGAGAGAGAAAGGACAAAGAGCGAGAAAGTAGACGAGGAGAAAG 1016
Db 175044 CTGGAATCTCCGAGAGAGAAAGGACAAAGAGCGAGAAAGTAGACGAGGAGAAAG 175103
QY 1017 GAAAGACACAGACCGCTCGAGGAGCAGGAGCGCGCAGAGACACAGACGAGCGCCAGG 1076
Db 175104 GAAAGACACAGACCGCTCGAGGAGCAGGAGCGCGCAGAGACACAGACGAGCGCCAGG 175163
QY 1077 TTCTCAGCAGACTACAGAGAGTGGAGTGCCTTTTGAAGACAGAGTTGGTGACCTGGAG 1136
Db 175164 TTCTCAGCAGACTACAGAGAGTGGAGTGCCTTTTGAAGACAGAGTTGGTGACCTGGAG 175223
QY 1137 GCATCGTCAGAGAGAGTGTGCTCCTTTGGCAACCGAAGTGTTCATGAGAGATGGAA 1196
Db 175224 GCATCGTCAGAGAGAGTGTGCTCCTTTGGCAACCGAAGTGTTCATGAGAGATGGAA 175283
QY 1197 GCCCAACAAGAGTGTGTCAGAGAGTCCACGTGAGACACCGTGCAGAGAGACAGAGGAGGAG 1256
Db 175284 GCCCAACAAGAGTGTGTCAGAGAGTCCACGTGAGACACCGTGCAGAGAGACAGAGGAGGAG 175343
QY 1257 CAGGAGGAGGAGGAGAGCTGAAGGGGCGTGGTGAAGAGAAACAGGAGAAATCCTTG 1316
Db 175344 CAGGAGGAGGAGGAGAGCTGAAGGGGCGTGGTGAAGAGAAACAGGAGAAATCCTTG 175403
QY 1317 CCCCTGAGAAACTGCTGAGCCCGAGGAGGTCCCCCAGGAAGCTGAGCCTGCTGAGGAG 1376
Db 175404 CCCCTGAGAAACTGCTGAGCCCGAGGAGGTCCCCCAGGAAGCTGAGCCTGCTGAGGAG 175463
QY 1377 CTGATGAAGAGCAGAGAGATGTGTCTCTGAGAGGAGACACACTCAACTGACAGACCTA 1436
Db 175464 CTGATGAAGAGCAGAGAGATGTGTCTCTGAGAGGAGACACACTCAACTGACAGACCTA 175523
QY 1437 AGTCTCTGAAGAGAGACGCTGCTGCCAAACACCCAGAAAGGCAATTCAGTGAAGTGGAGATG 1496
Db 175524 AGTCTCTGAAGAGAGACGCTGCTGCCAAACACCCAGAAAGGCAATTCAGTGAAGTGGAGATG 175583
QY 1497 CTGCTCTCTCAGAGAAAGATCAAGGTACAGGGAAGTCCCTTTGAAGAAACCTCTTCAGTAGC 1556
Db 175584 CTGCTCTCTCAGAGAAAGATCAAGGTACAGGGAAGTCCCTTTGAAGAAACCTCTTCAGTAGC 175643
QY 1557 TCAGGCTTAAAGAGCTGTCTGGGAGAGACAGAGGGGAAACGAGAGGTGGGGAGAC 1616
Db 175644 TCAGGCTTAAAGAGCTGTCTGGGAGAGACAGAGGGGAAACGAGAGGTGGGGAGAC 175703
QY 1617 GAAGAGCTGGAGAAATACCAACATTCACACCGAATCCCCAGAGAGTGTGATGAGCAG 1676
Db 175704 GAAGAGCTGGAGAAATACCAACATTCACACCGAATCCCCAGAGAGTGTGATGAGCAG 175763
QY 1677 AAGGAGAGAGCTCTGCTGCTGCTCCCGAGGAGCCTGAGAGACACAGCTGTCTGGAGAAA 1736
Db 175764 AAGGAGAGAGCTCTGCTGCTGCTCCCGAGGAGCCTGAGAGACACAGCTGTCTGGAGAAA 175823
QY 1737 GGGCCGCTGAGAGCACCCAGGAGTGGGAGCTGAGAGAGGAACTACTTCCGATGGAGAG 1796
Db 175824 GGGCCGCTGAGAGCACCCAGGAGTGGGAGCTGAGAGAGGAACTACTTCCGATGGAGAG 175883
QY 1797 AAGAGAGAGAGGGGATCACTCCCTGGGCATCCTTCAAAAAGATGGTGACACCCAGAGAA 1856
Db 175884 AAGAGAGAGAGGGGATCACTCCCTGGGCATCCTTCAAAAAGATGGTGACACCCAGAGAA 175943
QY 1857 CGGGTCCGAGAGCTTCTGAGAGTGAAGAGGAGAAAGAGCTGGAGAAAGGTCAAGAGCGCC 1916
Db 175944 CGGGTCCGAGAGCTTCTGAGAGTGAAGAGGAGAAAGAGCTGGAGAAAGGTCAAGAGCGCC 176003
QY 1917 ACCTTCTCTCCACTGATAGCAGTGTGAGAAATGCAAGATCAAGTCAAACTGTGGT 1976
Db 176004 ACCTTCTCTCCACTGATAGCAGTGTGAGAAATGCAAGATCAAGTCAAACTGTGGT 176063
QY 1977 GAGGAAACAAAGCCAGAGGAAACCAAGCGTAGGGTGGATCTTCAAGTGTCTGGGAGCA 2036
Db 176064 GAGGAAACAAAGCCAGAGGAAACCAAGCGTAGGGTGGATCTTCAAGTGTCTGGGAGCA 176123
QY 2037 CTGATTTGTTCGGATCATCAAGAGAGAGCAAGGAGGCAATCTCTTTCAGATGATGA 2096
Db 176124 CTGATTTGTTCGGATCATCAAGAGAGAGCAAGGAGGCAATCTCTTTCAGATGATGA 176183
QY 2097 GAGGGCCCAAGAGACACTGGGAGGGGACAGTCAAGAGAGAGGAGGCCAGCAAGACAA 2156
Db 176184 GAGGGCCCAAGAGACACTGGGAGGGGACAGTCAAGAGAGAGGAGGCCAGCAAGACAA 176243
```


QY 2157 GAAGCCGGAACAGACGCTGTTCTTGCCAGCACCCAGAGCGAGCAACAGCCGAAGAACT 2216
Db 176244 GAAGCCGGAACAGACGCTGTTCTTGCCAGCACCCAGAGCGAGCAACAGCCGAAGAACT 176303
QY 2217 TCCCTACCCGAGCCAGGGGGAAGGCCCTTCCGAAGGGGGAAGGTGCTCCACTTGGGAGTCA 2276
Db 176304 TCCCTACCCGAGCCAGGGGGAAGGCCCTTCCGAAGGGGGAAGGTGCTCCACTTGGGAGTCA 176363
QY 2277 TTTAAAGATTAAGTCACTCCAAAGAAAAATCCAAAGTCAAACTGGAAAGAAAGCCGAA 2336
Db 176364 TTTAAAGATTAAGTCACTCCAAAGAAAAATCCAAAGTCAAACTGGAAAGAAAGCCGAA 176423
QY 2337 GACTCTAGTGAAGCACTGTGCTCACTGAGATCCAACTGAGTGAAGAAAGATCTTGGGTT 2396
Db 176424 GACTCTAGTGAAGCACTGTGCTCACTGAGATCCAACTGAGTGAAGAAAGATCTTGGGTT 176483
QY 2397 TCCATTAAAGAAATTCATCCCGGAGCGGGGAAAGGAGGCGAGCGGAGCAAGAA 2456
Db 176484 TCCATTAAAGAAATTCATCCCGGAGCGGGGAAAGGAGGCGAGCGGAGCAAGAA 176543
QY 2457 GCCACTGGAAGAACTCAGGGCCAGTGAAGATTAATGAAGAGCACTTAATGTCCAGCC 2516
Db 176544 GCCACTGGAAGAACTCAGGGCCAGTGAAGATTAATGAAGAGCACTTAATGTCCAGCC 176603
QY 2517 GTCTGCTCTGTCTGAGTATATGCACTGAGAGGAGAGAGATGAAAGCCAGGGGAAAT 2576
Db 176604 GTCTGCTCTGTCTGAGTATATGCACTGAGAGGAGAGAGATGAAAGCCAGGGGAAAT 176663
QY 2577 ACGAGCTGCCCCAGCTCTGGGGGCTGTGTAAGTCCGAGAGACTCAATGAAGTCTG 2636
Db 176664 ACGAGCTGCCCCAGCTCTGGGGGCTGTGTAAGTCCGAGAGACTCAATGAAGTCTG 176723
QY 2637 GTCCACATGTAAGTGTGCACTGATGAGACCAAGGCGAGTCAACAGTGTCAAGAG 2696
Db 176724 GTCCACATGTAAGTGTGCACTGATGAGACCAAGGCGAGTCAACAGTGTCAAGAG 176783
QY 2697 CGCTGCTCTGTGAGTATCCGCTTCCGTAACAAGAACTCTTGAACACACAGCCGAGAA 2756
Db 176784 CGCTGCTCTGTGAGTATCCGCTTCCGTAACAAGAACTCTTGAACACACAGCCGAGAA 176843
QY 2757 GCCATGCCACCTGTGAAGAGTCACTGAAAAAGACATCTGCAAGAAAGCTCCTGTG 2816
Db 176844 GCCATGCCACCTGTGAAGAGTCACTGAAAAAGACATCTGCAAGAAAGCTCCTGTG 176903
QY 2817 CTCACCCAGACCTTACAGAGGGTAAAGATGCCATGACAGATGTGCAACAAGTGAAGT 2876
Db 176904 CTCACCCAGACCTTACAGAGGGTAAAGATGCCATGACAGATGTGCAACAAGTGAAGT 176963
QY 2877 GATTTCACCTCAGAAAGCTGTGACAGCCACAGAGACTTCAAGAGCTTCCGTACTGAAGA 2936
Db 176964 GATTTCACCTCAGAAAGCTGTGACAGCCACAGAGACTTCAAGAGCTTCCGTACTGAAGA 177023
QY 2937 GTTACCCGAAGCATTCGGGGGCCGAAGAGACCAAGACATGCTGTCCGACATTTCCAGCTG 2996
Db 177024 GTTACCCGAAGCATTCGGGGGCCGAAGAGACCAAGACATGCTGTCCGACATTTCCAGCTG 177083
QY 2997 ACTGACTCCCGAGACCAACAGAGAAAGCCACCCAGTTCAAGAGGTGAAGATGTGTG 3056
Db 177084 ACTGACTCCCGAGACCAACAGAGAAAGCCACCCAGTTCAAGAGGTGAAGATGTGTG 177143
QY 3057 CTAGATACAGAAAGAGAGCGCCAGAGCGAGGCCATCTTCCAGCCGTTGCAAGAA 3116
Db 177144 CTAGATACAGAAAGAGAGCGCCAGAGCGAGGCCATCTTCCAGCCGTTGCAAGAA 177203
QY 3117 GTGAAAAGAGAGTCCCAAGTGTCTGCAACCCAGACTGTGCAAGAAAGCGGGTCAAAAGCA 3176
Db 177204 GTGAAAAGAGAGTCCCAAGTGTCTGCAACCCAGACTGTGCAAGAAAGCGGGTCAAAAGCA 177263
QY 3177 CTGGAAGAGTTGAGAGGTGAAGAGAGACTCCGAAGTGTGCTTCCGAGAAAGAGAA 3236
Db 177264 CTGGAAGAGTTGAGAGGTGAAGAGAGACTCCGAAGTGTGCTTCCGAGAAAGAGAA 177323

QY 3237 GACGTTATGCCGAAGAGACCCGTGACAGAAAGCTGAGCTGAGCATCTTGACACAGGGCTT 3296
Db 177324 GACGTTATGCCGAAGAGACCCGTGACAGAAAGCTGAGCTGAGCATCTTGACACAGGGCTT 177383
QY 3297 GAGACTGGAACAGCTACTCAGAGAGCTTGAAGTCTTGAAGTCAACGCGATGTAGAC 3356
Db 177384 GAGACTGGAACAGCTACTCAGAGAGCTTGAAGTCTTGAAGTCAACGCGATGTAGAC 177443
QY 3357 CAGTGGCCAGCGGCGAGGTTATCAAGCTCCAGAGCTGATGGAACAGGGCCGCGCCCT 3416
Db 177444 CAGTGGCCAGCGGCGAGGTTATCAAGCTCCAGAGCTGATGGAACAGGGCCGCGCCCT 177503
QY 3417 GAGTCATCCGAACCTTGAAGACAGTGAACAATGAAAGCACTCCCTTGAAGCAATTA 3476
Db 177504 GAGTCATCCGAACCTTGAAGACAGTGAACAATGAAAGCACTCCCTTGAAGCAATTA 177563
QY 3477 GACACTGCAATGGGACACAGCAAGATGAAACATTTGACAGCCAGACAGTAAAGCACT 3536
Db 177564 GACACTGCAATGGGACACAGCAAGATGAAACATTTGACAGCCAGACAGTAAAGCACT 177623
QY 3537 GCAGCTGCAAGGCACTCAAGTGTCAAGAAAGAGCGGCTACTGCTCAAGAAAGAGAG 3596
Db 177624 GCAGCTGCAAGGCACTCAAGTGTCAAGAAAGAGCGGCTACTGCTCAAGAAAGAGAG 177683
QY 3597 CCTTCGACACTAACCTTAATTAATGTTCCAGCCAGGAAGAACATGGGGAAGAACAGAA 3656
Db 177684 CCTTCGACACTAACCTTAATTAATGTTCCAGCCAGGAAGAACATGGGGAAGAACAGAA 177743
QY 3657 GATGTTCTTGAACCTTACACAGCAAGAGTTACTGCTGACGCTGCTGCTTGGCAAG 3716
Db 177744 GATGTTCTTGAACCTTACACAGCAAGAGTTACTGCTGACGCTGCTGCTTGGCAAG 177803
QY 3717 ACTGAGGTGGGTCAAGAGGTGAGTGTGATGTTGATGGAAGAAAGTCAAAAGAA 3776
Db 177804 ACTGAGGTGGGTCAAGAGGTGAGTGTGATGTTGATGGAAGAAAGTCAAAAGAA 177863
QY 3777 CAGAGAGTGTGTAACACTGTGCAACCAAGTCAAAAGCTGTGATGTGACATATGAC 3836
Db 177864 CAGAGAGTGTGTAACACTGTGCAACCAAGTCAAAAGCTGTGATGTGACATATGAC 177923
QY 3837 AGTGAAGTGAATGGAGTGGCCGGGTGTCAAGAAAGAGAGTACTGAAGTCAAGTCTT 3896
Db 177924 AGTGAAGTGAATGGAGTGGCCGGGTGTCAAGAAAGAGAGTACTGAAGTCAAGTCTT 177983
QY 3897 AGCTGGAAGAGGAGAGATGGAATGTAACGTTGAAAGGAGAAAGGAGAGCAAAAGCA 3956
Db 177984 AGCTGGAAGAGGAGAGATGGAATGTAACGTTGAAAGGAGAAAGGAGAGCAAAAGCA 178043
QY 3957 GAGCAAGTGAATGAAGAGTGAAGAGAAACAGCCGCTCTGAGCATGAAGAACTTAC 4016
Db 178044 GAGCAAGTGAATGAAGAGTGAAGAGAAACAGCCGCTCTGAGCATGAAGAACTTAC 178103
QY 4017 GGGAAAGCAATCTGACACTTGAATGCTCCAGCTCAGAGAGGGGAGAGCACTGGAAAGC 4076
Db 178104 GGGAAAGCAATCTGACACTTGAATGCTCCAGCTCAGAGAGGGGAGAGCACTGGAAAGC 178163
QY 4077 CTGGAAGAGAGCTTCTCTCCCAAGCAACAAAGAGTGTGATGAGTGTCAAGT 4136
Db 178164 CTGGAAGAGAGCTTCTCTCCCAAGCAACAAAGAGTGTGATGAGTGTCAAGT 178223
QY 4137 CAAGACCTGGAACACAGTCACTCAACAGCAGAAAGCTGTGAAAGAGTCAATGAAGC 4196
Db 178224 CAAGACCTGGAACACAGTCACTCAACAGCAGAAAGCTGTGAAAGAGTCAATGAAGC 178283
QY 4197 GTTGTGATTTCAAGACAGGTGAAGTCCAGATGTGATGACACTTATTAACAGCT 4256
Db 178284 GTTGTGATTTCAAGACAGGTGAAGTCCAGATGTGATGACACTTATTAACAGCT 178343
QY 4257 GAGAAAGTCTCTGCAACGGGTGCGCACTGTGACTTTCAGCATGCAAGAGCAACGGTACC 4316
Db 178344 GAGAAAGTCTCTGCAACGGGTGCGCACTGTGACTTTCAGCATGCAAGAGCAACGGTACC 178403
QY 4317 CTGGGGCTGAGTCTCAGGCAAGATCCCAATCATAGTAATCTTGCCTCTGAAGAC 4376

Oy	898	CTCAGGTTGGGCCCGGCTGGCGCAGAAAGACCAAGCTTCAAGAAATCAAAAGAGATGATC	957
Db	707	CTCAGGGGTGGGCCCGGCTGGCGCAGAAAGACCAAGCTTCAAGAAATCAAAAGAGATGATC	766
Oy	958	TGGAAACTGCCGAGAAAGAAAGAGCAAGAGCGCAGAAAAAGTAAACGAGAAAGAAAAAG	1017
Db	767	TGGAAACTTCCAGAAAGAGAAAGAGCAAGAGCGTGAAMAAAGTAAACGAGAAAGAAAAAG	826
Oy	1018	AAAAACAGAGCCAGCC---TCGAGAGACACAGAGCCGACAGAAAGACACAGACCAAGGCCA	1074
Db	827	AAAAACAGAGCCAGCC---TCGAGAGACACAGAGCCGACAGAAAGACACAGACCAAGGCCA	886
Oy	1075	GCTTGTCAAGCAGACTACGAGAAAGTGGAGCTGCTTTGGAAAGCAACAGTTGGTGAACCTGG	1134
Db	887	GCTTGTCAAGCCAGCTATGTAGAAAGTGGAGTTCCTTTGGAAAGCCAGGTCCGGTGAACCTGG	946
Oy	1135	AGGCATCGTCAAGAGAGAACTGTGCTCTTTGGCAACGGAAGTGTGATGTAGGAAGATG	1194
Db	947	AGGCATTTGTC---GGAGAGTGTGCTCTTTGGCAACGGAAGTGTGATGTAGGAABACGG	1003
Oy	1195	AAGCCACCAAGAAAGTGTGTGACAGAGTCCACGTGAGCAACCGTGGAGAAACAGAGAG	1254
Db	1004	AAGCCACCAAGAAAGTGTGTGACAGAGTCCACGTGAGCAACCGTGGAGAAAGTACAGAAAG	1063
Oy	1255	AGCAGGAGAGAGAGAGAGCTGAAGGAGGAGCGTGGTGGTAGAGAAACAGAGAAATCCT	1314
Db	1064	GGCA--AGAGAGAGCAGAGAGTGAAGGAGGAGTGTGTGTGGAGAGATGGGAGAAATCCT	1120
Oy	1315	TGCCCCCTAGAAACCTGGCTGAGCCCCCAGAGAGTCCCCCAGAAAGCTGAGCCCTGTAG	1374
Db	1121	TGCCCCCTAGAAACCTGGCTGAGACCAGAGAGTCCCCCAGAAAGCTGAGCCCTGTAGAG	1180
Oy	1375	AGCTGATGAAGAGCAGAGAGATGTGTGCTTCTGAGAGAGACCACTCAACTGACAGACC	1434
Db	1181	AGCTGATGAAGACCAAGAAAGTATGTGCTCTGTGGGATGACCAATCTCAGCTGACAGATC	1240
Oy	1435	TAACTCTTGAAGAGAGACGCTTGCCCAACACCCACAAAGGCATTGTCAAGTAGAGTGAAG	1494
Db	1241	TAACTCTTGAAGAGAGATGTCAACCAAAACCCCGAAGGCATTGTCAAGTAGAGTGAAG	1300
Oy	1495	TGCTGTCTCTCAGAAAGAAATCAAGGTACAGGGAAGTCCCTTGAAGAAATCTTCAAGTA	1554
Db	1301	TGCTGTCTCTCAGAGAGAAATCAAGGTACAGGGAAGTCCCTCGAAGAACTTTCAAGCA	1360
Oy	1555	GCTCAGGCTTAAAGAGAGTCTGTGGAGAGAGCAAGAGGAG---ACGAGAGGTGGGG	1611
Db	1361	GTTTCGGGCTTAAAGAGAGTCTTCGGGAAAGAGCAAGAGGAGAGAGAGAGGAGGG	1420
Oy	1612	GAGACGAAGGCGCTGAGAAATACCAACACTTCAACCGAATCCCCGAGAGAGTGTATG	1671
Db	1421	GAGATGAAGGCCAGAGGAATACCAACACTTCAAAACCGAGTCCCGAGAGAGTGTACG	1480
Oy	1672	AGCAGAAAGGAGAGAGCTGTGCGTGTCTCCCGAGAGAGCTTGAAGAGACCAAGTGTCTGG	1731
Db	1481	AGCAGAAAGGAGAGAGCTGTGCGTCTTCCCTGAAAGAGCCCGAGAGAGATCGCGTGTCTGG	1540
Oy	1732	AGAAAGGCGCTGAGAGACCCCAAGATGGGAGAGCTGAGAGAAAGAACTAACTTCCGATG	1791
Db	1541	AGAAAGGCGCATCGGAAGCAACCCAGGAACGGAAACCTGAGAGAGGAGCACTTCCAGC	1600
Oy	1792	GAGAGAAAGAGAGAGAGGATCACTCCCTGGGCACTCTTCAAAAAGATGGTGACACCA	1851
Db	1601	GAGAGAAAGAAAGAGAGGATCACTCCCTGGGCACTCTTCAAAAAGATGGTGACACCA	1660
Oy	1852	AGAAACGGGTCGAAAGACTTCTGAAGTGAACAAGAGAGAAAGACTGAGAGAGTCAAGA	1911
Db	1661	AGAAACGGGTCGGAAGACTTCTGAAGCGCAACAAGAAAGAGCTGGAATGAGGTCAAGA	1720
Oy	1912	GCGCCACTTGTCTCTCACTGATAGCAACAGTTCAGAAATGCAAGATGAAGTCAAAACTG	1971
Db	1721	GTCACACTTGTCTCTCACTGAGAGCAACGGGTGTGAATGCAAGATGAAGTTCAGAGCGG	1780

QY	1972	TTGTGTAGAGAACAAACCGCAGAGAACCAAGCGTACGGTGGATCTTCACTGTCCTTGGG	2031
Db	1781	TTGGCGAGGAGCAAAAGGTCAAGAGAGCCAAAGCGCAGGGTGATCTTCACTGTCCTTGGG	1840
QY	2032	AAGCACTGATTTGTGTGGGATCATCCCAAGAAAGAGAGCAAGAAAGCATCTCTTCATATG	2091
Db	1841	AGGCGTTGATTTGTGTGGATCTGTCCAGAAAGAGAGCGAGAGAGCATCTCTTCAGATG	1900
QY	2092	ATGAAGAGGGCCAGAGCACTGGGAGGGGACAGTCACAAGCAGAGAGGCGCAGCAAG	2151
Db	1901	ATGAAGAGGGCCAAAGAACACTGGGAGGGGATGGCCACAGAGCGAGAGGCTAGCAAG	1960
QY	2152	ACAAAGAGCCGGAAACAGACGCTGTTCTGTCCAGCACCAGAGACAGACCAAGCCGAG	2211
Db	1961	ACAAAGAG-----CAGATGCTCTTCTGTCCAGCACCAGAGAACAGACCAAGCCGACG	2014
QY	2212	GAAGTTCTCCACCCGAGCAGCGGGAAGCCCTTCCGAAGGGGAAGTGTCCTCACTTGGG	2271
Db	2015	GAAGTTCTTCACCCGAGCCAGCTGGAAAGCCTTCTGAAGGGGAGGGCGTCTCCACTGGG	2074
QY	2272	AGTCATTTAAAAAGATTAGTCACTCCAAAGAAAAAATCCAACTCAAACTGGAGAGAAAG	2331
Db	2075	AGTCATTTAAGAGATTAGTCACTCCACAAAAAATCAATCAAACTGGAGAGAGAG	2134
QY	2332	CCGAAGACTCTAGTGTAGAGCAGATTGTCCTAGATTCGAACCGAGATGAGAAATCTT	2391
Db	2135	CCGAAGACTCTCGGTGCAAGAGAGTGGCGCTCCGAGATCGAACCAAGATGAGAGAAATCTT	2194
QY	2392	GGGTTTCAATTAGAAATTCATCCCGGAGCGGCGAAGAAAGGCGAGACGGAGAGCAAG	2451
Db	2195	GGGTTTCAATTAGAAATTTATTCTTGGACGCGGAGAAAGGCGAGATGGAAAGCAAG	2254
QY	2452	AACAAGCACTGTGGAGACTCAAGGCGCAGTGGAGATTAATGAGACGACCTTAATGTC	2511
Db	2255	AACAAGCGCGCGTTGAAGACTCGGGGCCAGAGAGATCAATGAGAGACGACCCGACGTCC	2314
QY	2512	CAGCGTCGTGCTCTGTCTGAGATTAATGCACTGTGAGAGGGAAGATGGA-----	2563
Db	2315	CAGCTGTGTGTGCTCTGTCTGAGTACGATGTGGGTAGAGAGAAAGCTGGAAAGCGCAGC	2374
QY	2564	-AGCCACGGGGGAATCGAGAGCGCCCGACGTGCGGGGGCGTGTGATCGTGTCCAGAGAGC	2622
Db	2375	GAGCTCAGAGAGAACTGTGAGCGTCCCGACGTGAAGGGGGCTGTGTATGTGTCTGAGAGC	2434
QY	2623	TGAGTAAAGCTGTGTGCACACTGTGAGTGTGCGAGTCATTGATGGGACCAAGGCGATCA	2682
Db	2435	TTAGTAAAGCTGTGTTCACACTGTGAGTGTGCGGTCAATTGATGGGACCAAGGCGATCA	2494
QY	2683	CGAGTGTGAGAGAGCGGTCTCTTGTGTGATATCCGCTTCCGTAAACAGAACTCTTGAAC	2742
Db	2495	CGAGTGTGAGAGAGCGGTCTCTTGTGTGATATCTGCTTCATGACAAAGCCTCTTGGAGC	2554
QY	2743	ACAAGCCGGGGAAGACCAATGCCACTGTGTGAAGAGGTCACTGAAAAAGACATCATTTGAG	2802
Db	2555	ACGCAAGGGAAGTGTGCACACCGCTGTGTGAGAGTCACTGAAAAAGACATCACTGAG	2614
QY	2803	AAGAAACTCTGTGCTCACCCAGACGTTAACAGGGTAAAGATGCCATGACGACATGG	2862
Db	2615	AAGCAACTCTGTCACTCGCCGACGACTTTAACAGGGGGCAAAAGATGCCATGACGACATAG	2674
QY	2863	TCAACAGTAAGTGATTTCACTTACGAAGCTGTGACAGCCACAGAGACTTCAGAGGCTC	2922
Db	2735	TTCGCGCTTAAGAACTTTACCGAAGATCACTGGGGCAAGAAAGCCACAGACATGTGTCTG	2794
QY	2923	TTCGTACTGAAGATTAACCGAAGCATCGGGGGCCGAAGAGCCACAGACATGTGTCTCG	2982
Db	2795	CAATTTCCAGCTGTGACTGCCACAGACACCAAGAGAAAGCCACCCAGTTCCAGAGG	3042
QY	3043	TAGAGAGTGTGTGATTAGTACGAAGAAAGAGAGCGCCAGACGAGCCATCTTCAAG	3102

||||| TAGAGGTGGCATCTAGATACGGAAGACAGGAGCCCGACGCGCGCTCTCCAG 2914
Db
3103 CCCTTGACAGCAAGGTGAAGAGAGTCCAGAGTGCTGCAACCCAGACTGTGCAGAGAA 3162
Qy
2915 CCCTTGACAGCAAGGTGAAGAGAGTCCAGAGTGCTGCAACCCAGACTGTGCAGAGAG 2974
Db
3163 CGGGTCAAAAGCACTCGAGAGGTTGAGGAGGTAGAGGAGGACTCCGAAAGTGTGCTT 3222
Qy
2975 CAGGCCGAAAGCACTCGAGAGGTTGAGGAGGTAGAGGAGGACTCCGAGGTTGCTGCTA 3034
Db
3223 CGGAGAAAGAGAGGACCTTATGCGGAAAGACCCGCTGCAGGAGCTGGAGCTGAGCATC 3282
Qy
3035 CCAGAGAGAGAGGATGTTGTCGCGAAGGACCCGCTGCAGGAGCTGAACTGAGGATC 3094
Db
3283 TTGACAGAGGCTTGAGACTGGAACAGGCTACTCCAGAGAGCCCTTGAAGTTCTTGAAGTCA 3342
Qy
3095 TTGCACAGGGCTCCGAGACTGTACAGGCTACCCAGAGAGCCCTTGAAGTTCTTGAAGTCA 3154
Db
3343 CGGCAGATGTAGACCATGTGCGCACGTGCCAGGTATCAAGCTCCAGCAGCTGATGGAAC 3402
Qy
3155 CAGAGGATGTAGACCGTGCACACATGCCAGGTATCAAGCACCCAGCAGCTGATGGAAC 3214
Db
3403 AGGCGGTGGCCCTGAGTCACTCGAAACCTTTGACAGACAGTGAGACAAATGGAAGCACTC 3462
Qy
3215 AGGCTGTGGCCCTGAGTCACTGNAACCTTTGACAGACAGTGAGACAAATGGAAGTACTC 3274
Db
3463 CTTAGCAGATTACAGACACTGCGAGATGGGACACAGCAAGATGAACCAATTTGACAGCCAGG 3522
Qy
3275 CCCTCGCAGATTACAGACACTCCAAACGGGACACAGCAAGACGAGACCGTTGACAGCCAGG 3334
Db
3523 ACAGTAAGCCACTGACGTGTGAGGAGTCAAGGTGACAGAGAGAGCGGCTACTG 3582
Qy
3335 ACAGTAATGTCATTGCGCCGCTCAAGCAGTCAAGGTCACTGAGAGAGGAGCGAGTCTG 3394
Db
3583 CTGAGAAAGAGGAGCCCTTGACACTACTAATATGTTCCAGCCCGAGGAGAGACATGGGG 3642
Qy
3395 CTGAGCGAGGGGCTTCAACCATC---TAGTTTCCAGCCCGAGGAGAAACACAGGG 3451
Db
3643 AAGAACAGGAAGAGATGTTCTTGAACCTACACAGCAAGAGCTTACTGTCGAGCCGCTGC 3702
Qy
3452 AAAAAACAGGAAGGATGTTCTAGAACCCACA---CAAGCGCTGGCTGCCGGGCGAGTGC 3508
Db
3703 CCCTTGCGCAAGACTGAGGTGGGTCAAGAGGTGAGGTGACTGTTGGATGGAGAAA 3762
Qy
3509 CTATTCTGGCAAGGCTGAGGTGGGTCAAGAGGTGAGGTGGCCAGTTTGTATGGAGAAA 3568
Db
3763 AAGTCAAAG-----AAGAACAGGAGGTGTTGTACACTCTGGACCCCAACA 3807
Qy
3569 AAGTCAAAGACGACAGTGTGTTAAAGAACTGGAGGTGCTGTGCACACTGGACCCCAACA 3628
Db
3808 GTCAAAAGGCTGTGATGTGACATATACACAGTGAAGTGAATGGAGTGGCCGGGTGTCCAGG 3867
Qy
3629 GTCAAAAGACTGTGACTTGACACGTGACAGTGAAGTGAATGGAGTGGCCAGATGTCCAGG 3688
Db
3868 AAAAGAGAGTACTGAAGTGCAGGTCTTACCTGGAGGAGGAGAGATGGAAACTGACG 3927
Qy
3689 AAATCAGAGTAATGAAGAAACAGAGTATTAGCCCGGAGAAAAGAGAGATGGGAACCGACG 3748
Db
3928 TTCAAAAGGAGAAAGGGACAAAGCCAGACAGTGAAGTGAAGAGGTGAGCAGGAAA 3987
Qy
-3749 TTGAAGAGGAGGAAAACAGAGCAAGAACAGACAGCAAGCCAGTGAAGAAACATGAGCAGGAAA 3808
Db
3988 CAGCCGCTCTGAGCATGAAGGAACCTTACGGGAAGCCAGTCTTGACACTTGAATGCCCCA 4047
Qy
3809 CAGCTGCTCTGAGCATGAAGGAACCCACCCCTAAGCCAGTCTTGACAGCTGACATGCCCTC 3868
Db
4048 GCTCAGAGGGGAAAGGCACTGGGAAGCCCTTGGAGGAAGCCCTTCTCTCCAGACCAAG 4107
Qy
3869 ACTCAGAGGGGAAAGGCACTGGGAGCCCTTGAAGGAAGCCCTTCTCTCCAGACCAAG 3928
Db
4108 ACAAGCAGGTTGCATAGAGGTTCAAGTTCAAAGCCTGGACACACAGTCACTCAAAACAG 4167
Qy
|||||

3929 ACAAAGCAGATTGCATAGAGGTTCAAGTTCAAAGCTCAGACACACAGTCACTCAAAACA 3988
Db
4168 CAGAAGCTGTGAAAAAGGTCTATAGAAAACGGTTGTGTGATTTTCAGAGACAGGTGAAAGTCCAG 4227
Qy
3989 CCGAAGCTGTGAAAAAGGTTCGAAGAAAACGTGTGGCAACTTCAGAGATGGATGAAAGTTGG 4048
Db
4228 AGTGTGTAGGTGCACACTTATTACACGAGCTGAGAGTCTCTGCAACGGGTGGCCACTGGA 4287
Qy
4049 AGTGTGCAGGTGGCAATCATTTACCAGCTGAGAAAGCTCTCCGAAACCGGTGGCTACGGGA 4108
Db
4288 CTCTTCAGCATGACAGAGGACACGGTATACCTGGGGCTGAGTCTCAGGACAGATCCATCC 4347
Qy
4109 CTCTTCAGCATGAGAGAGGACACCGTGCAGGAGGCTGAGTCTCAGGACAGATCCATCC 4168
Db
4348 CAATCATAGTAACTCTCTGCTCTCTGAAAGCAACCTACATCTGACCTTCAAGAGGAATAA 4407
Qy
4169 CAATAATAGTAACTCTCTGCTCTCTGAAAGCATCTCATTTCTGACCTTCAAAGAGAGTGA 4228
Db
4408 GGCATCCACAGAGAGCGATCAGAGAGAGAGGACAAAGCCAGATGCTGTCTCTGATGCTG 4467
Qy
4229 GGCATCCACAGAAACAGAGATCAGATGAAGATAAAGCCAGATGCTGTGCTCTGATGCTG 4288
Db
4468 ACGCAAGGAGAGTACAGCAATCGAAAAAGTCTCAAGGCTGAACCTGAGATCCTGGAAC 4527
Qy
4289 CCGCAAGGAGAGTGCAGACAGAGAGAAATCTCAGGCTGAACCTGAGATCTTGGAAAC 4348
Db
4528 TTGAGAGTAAAGAGCAACAAGATTGTCTGAAACGTCAATTACAGACAGCCGTTGACAGTTG 4587
Qy
4349 TTGAGAGTAAAGAGCAATAAGATTGTCCAGAGTGTCTATCCAGACAGCGCTCGACAGTTG 4408
Db
4588 CAGTACAGAAACAGCCCGGAAACCTCATGCTTTATGATTTCAAGACCCAGGTTCTGTG-CA 4646
Qy
4409 CAGTACAGAAACAGCCCGGAAACCCAGCTTCTGATTTACAGAAATCAGGTTCTCTGTGA 4468
Db
4647 TGCAAGCTTTGACAGAGGAGGCCCAACAGATGCTGCACAAAATGAAGATG----- 4698
Qy
4469 TGCAGGCTGACAGCCAGGAGGACAAACAGATGCTGCAAGAGTGAAGGACGCTTCAAG 4528
Db
4699 -----CCAAGATGAACACACCCAGTGCAGCCAGCCAGAGAGACTTGAAGTCTCTGACC 4751
Qy
4529 TCTCCCCCAAGATGGAACACTCAGTGCCTGAGCCAGGAGGACTTGCCTGTTCTGATA 4588
Db
4752 GTTCTGAGAGCATGGCTCAGCCTCGGAAATGCTTCCGCGCTTGAGTTGAAAGCGCCG 4811
Qy
4589 GTTCTGAAGGATGAGCAAGGCTTCAGAAATGATCAACCGCTTGACAGTTGAAAGTGCCA 4648
Db
4812 GTGCTAAAGTAAAGCATTTGAGAAAGTCCCTCCTCAACCCCAAGATCCAAAGGAGCATGCT 4871
Qy
4649 GTGCTAAAGAAAGTGTGAGAAAGTCTCCTCTTCAGTGCAAAAGAT-GAAAAGGAAACATGCT 4707
Db
4872 GCTGATGGCCCTCAGCTCCAAAGCTTAGCCAGGACAGGCGCAGTGCCTCTGGAAACCTTA 4931
Qy
4708 GCTGACGGCCCCCAGCACCAAGCTTAGCCAGGACAGGCGGATGCCTCTGGAAATCTTA 4767
Db
4932 ACCAAAGAAATCCCCAGACACCAACGGACCAAAAGCTTAACCGAGGAGGCGATCCCCAAA 4991
Qy
4768 ACCAAAGAGTCCCCAGACACCAACGGACCAAAAGCTAAACCGAGGAGGAGATGCCCTGAA 4827
Db
4992 GTTGAAGTCCAGGAGAGAAATGTCTACCAAGTCAAGAGAAACAAGGCCCGAGGCA 5051
Qy
4828 GAAG-----NAATGAACAGGCCCGACGA 4851
Db
5052 GAAGAG---GACTGCGAGGAGCAAAAGGAGAGACCTGGCAGAAATCTTAAGATGTTAGTTGC 5108
Qy
4852 GAAGAGAGAGCACTACAGAGGCCAAAGGAGAGACCTGACAGAAATCTTAAGACGTGAGTTGC 4911
Db
5109 TCATTGTCATCTGTAAGACCAAGATGTGAAACCAAGTCAAGAAACAAGATGCTGCTGTT 5168
Qy
4912 TCAT-----TGTAAGAACTAGATGTGAAGTGAAGTCAAGAAACAAGATGCTGCTGTT 4963
Db
5169 GGGACCTTGAGACCAAGATTTCAAGCCCATGAGATCCAGAGAGGAGGCCCTGCAATGA 5228
Qy
4964 GGGACCTTGAGACCAAAATTTTCAAGCCCTTGAGGTGCGAGAGAGAGAGCCGTTCCAATGA 5023
Db
|||||

QY 231 CAAGGAGATGTCATGTCCAGAGAGAAACACGAGGG----- 268
Db 402 CAAGATGAGCTCAGCCTCCAGGAGGGTGACCTAAATGGCCAGAAAGAGCCCTGAACGGT 461
QY 269 -----GCAGGAGGAAGAGTCCGTGATGAGGATGTTGGACACGGA 308
Db 462 CAAGGAGCCCTAAACAGCCAGGAGGAAGAAGTCAATGTCACAGAGGTTGGACAGGA 521
QY 309 GAGTCAGAGATGTGAGAGAAAGACCGAGTTGAAGAAATGGCGGCAACTCCACAGCT 368
Db 522 GACTCTGAAGATGTGAGCGAAAGAGATCCCGATAAAGAGATGGCTACTAAAGTCAGCGGTT 581
QY 369 GTTGAAGATATCAGAAAGATGGCGAGGAGACATCAGAAATTAATTTGAACAGATCCCT 428
Db 582 GTTCAACACATCAGATGATGGCGAGGAGACACCCGAAATAATTCGAAACAGATTCCT 641
QY 429 GCTTCAGAAAACATGTGGGAAGAAATGGTACAGCCTCTGAGTCCCAAGGCTAATGATGTT 488
Db 642 TCTTCAGAAAGCAATTTAGAAGAGCTAACACACCCACTGAGTCCCAAGGCTAATGATTT 701
QY 489 GGCTTCAAGAAAGTATTTAAATTTGTTGTTTAAATTTCA CGTGAAGAGGATAAAAT 548
Db 702 GGATTTAAGAAAGTGTAAAGTTGTTGGCTTTAAATTTCACTGTGAATAAGGATAAGACA 761
QY 549 GAAAAGTCAGATGTCCTCACTACTCTCAAGAGGATGAAGCGAAGGGCGAGAA 608
Db 762 GAGAAGCCTGACACTGTCCTCACTACTCTCAAGAGGATGAAGGGGAGGGAGCGACA 821
QY 609 GCCTCTGCGGAGCTGGAGACCCAGGAGCCAGTGTGGAGACTGCCGTCCGAGAGTCA 668
Db 822 -----GGGCTGGGACCCAGGACCCAGCCTTGGGGCT-----GGAGAAGCA 866
QY 669 GCATCCAAAAGAAAGTGAGCTGAAGCAATCCACAGAGAAAGCAAGGACACCTTGAAGCAA 728
Db 867 GCATCCAAAAGAAAGCGAAACCCAAACATCTACAGAGAAACCCGAAGAGACCTTGAAGCGT 926
QY 729 GAACAGAGCAGCAGAGAAATCCCCCTTCAAGCCGAATCTGATCAAGCGGCTGAGGAAGAA 788
Db 927 GAGCAAGCCACGAGAAATTTCTCCCGAGCCGAATCTGGCCNAAGCA---GTGGAGGAA 983
QY 789 GCCAAGATGAAGGAGAAAGAAACAGAGAAAGAGCCACCAAGTCCCGCAGAAATCCCG 848
Db 984 TGCAAGAGAGAGAGAGAGAAACAAGAAAGAACTAGCAAGTCTGCAGAAATCTCCG 1043
QY 849 AGAGGCCAGTCAACAGTGAGACAACATCTTCTTCAAGAAAGTTCTTCACTCAAGGTTG 908
Db 1044 ACTAGTCCCGTGACAGTGAACAGGATCAACCTTCAAAAAATTTCTTCACTCAAGGTTG 1103
QY 909 GCGGCTGGCGCAGAGACAGCTTCAAGAAATCAAGAGAGATGATCTGAAAACCTGCC 968
Db 1104 GCGGCTGGCGCAAAAAGCACAGTTTCAGGAAGCCGAAGGAGATGAAGTGGAAAGCTTCA 1163
QY 969 GAGAAGAAAGGAGCAGAGGAGGAGAAAGTGAAGAGAGAAAGAAAGGAAAGACAGAG 1028
Db 1164 GAGAGNAAAGGACAAAGAGCCAGAAAGATAGACACAGAAAGACGGAAGGCGAGAG 1223
QY 1029 CCAGCCTCGGAGGA-----GCAGGAGCCGCGCAGAGAC 1061
Db 1224 GTTGGCTCCGAGAACTGACCGCTCCGAGCAAGCCACCCACAGGAGCCGCGAGAAAGT 1283
QY 1062 ACAGACCAGGCGAGTTGTGACAGAGCTACAGAGAGTGAGCTGCTTTGGAAAGACCAG 1121
Db 1284 GCCCA CGAGCCCGGTTATCAGCTGAATATGAGAAAGTTGAGCTGCCCTCAGAGGAGCAA 1343
QY 1122 GTTGTGACCTGGAGGATCGTCAGAGGAGAGTGCTCTTCTTGGCAAAGGAGTGT 1181
Db 1344 GTCAGTGGCTCGGAGGACCTTCTGAAGAGAAACCTGCTCCGTGGCGACAGAAAGTGT 1403
QY 1182 GATGAGAAGATGAAGCCCAACCAAGAA---GTTGTTGCAGAGGTCCACGTGAGCACCGTG 1238
Db 1404 GATGAGNAAATAGAGTCCACCAAGAGAGAGTTGTGGCCGAAAGTCCACGTGAGCACCGTG 1463
QY 1239 GAGAAGACAGAGGAGGAGGAGGAGAGAGGCTGAAGGGGGCGTGGTGTAGAA 1298

Db 1464 GAGGAGAGAAACCAAGACAGAAAA-----ACGGAGGTGGAA 1499
QY 1299 GGAACAGGAGAAATCCTTGGCCCCCTGAGAAACTGGCTGAGCCCCCAGAGGTCCCCCAGGAA 1358
Db 1500 GAAACAGCAGGGTCTGTGCCAGCTGAAGAAATGGTTGAATGATGCAGNAACCTCAGGAA 1559
QY 1359 GCTGAGCCTGCTGAGAGCTGATGAAGAGCAGAGAGATGTGTCTCTGGAGAGACCAC 1418
Db 1560 GCCGAACCTGCCAAGAGCTGTGAAGCTCAAAAGAAACGTGTGTTTCCGAGAGGACCT 1619
QY 1419 ACTCAACTGACAGACCTTAAGTCTCTGAAGAGAGACGCTGCCCAACACACAGAGGCAAT 1478
Db 1620 ACACGGAGCTGACCTCAGTCTCTGATGAAGGCTGTGTCCAAACCCCCCGAAGGCGTT 1679
QY 1479 GTCAGTGAGTGGAGATGCTGTCTCTCAGGAAAGAAATCAAGCTCAGGGAAGTCCCTTG 1538
Db 1680 GTGAGTGAGTGGAAATGCTGTCTATCAGGAGAGAAATGAAGTGCAGGGAAGTCCACTA 1739
QY 1539 AGAAAATCTTCAGTAGTCTAGGCTTAAAGAGCTGTCTGGGAAGAGCAAGAGGGGAAA 1598
Db 1740 AGAAGCTTTTACACAGCACTGGCTTAAAGAAAGCTTTCTGGAAAGAAACAGAAAGGAAA 1799
QY 1599 CGAGAGGTGGGGAGACGAGAGCTGGAGATACCAACACATTCACACCGAATCCCCA 1658
Db 1800 AGAGGA---GGAGGAGACGAGGAATCAGGGGAGCACACTCAGGTTCCAGCCGATTTCCG 1856
QY 1659 GAGAGTGCTGATGAGCAGAAAGGAGAGAGCTGTGCTGTGCTCCCGAGGAGCTGAGGAG 1718
Db 1857 GACAGCCAGGAGGAGCAAAAGGCGAGAGCTCTGCTCTATCTCCCTGAGGAGCCGAGGAG 1916
QY 1719 ACCACGTGTCTGGAGAAAGGGCGCTGGAGCAACCCAGGATGGGAAGCTGAGGAAGGA 1778
Db 1917 ATCAGCTGTCTGAAAAGGGCTTTAGCCGAGGTGACAGAGATGGGGAAGCTGAAGAAGGA 1976
QY 1779 ACTACTTCCGATGAGAGAAAGAGAGAGAGGATCACTCTCTGGGATCTCTTCAAAAAG 1838
Db 1977 GCTACTTCCGATGAGAGAAAAAGAGAGAGGTGTCACTCTCTGGGATCATTTCAAAAAG 2036
QY 1839 ATGGTCACACCCAAAGAAAACGGGTCCGAAGACCTTCTGAGAGTGCACAAAGGAGGAGAGCTG 1898
Db 2037 ATGGTGAAGCCCAAGAAAGCGTGTAGACGGCTTCGAAAGTATTAAGAAAGATGAGCTG 2096
QY 1899 GAGAAGTCAAGAGCCCACTTTGCTCTCCACTGATAGCACAGTGTGAGAAATGCAAGAT 1958
Db 2097 GACAAGGTCAAGAGCGCTACCTTTGCTTCCACCGAGAGCACAGCTCTGAAATGCAAGAA 2156
QY 1959 GAAGTCAAAAATGTTGCTGAGGAAACAAAAGCCAGAGAACCAAGCGTAGGTGGATACT 2018
Db 2157 GAAATCAAAAGGGAGCGTGGAAAGAGCCAAAGCCGAAAGAACCAAGCGCAAGGTGGATACC 2216
QY 2019 TCAGTGTCTTGGGAAGCACTGATTTGCTGCGATCATCCAAGAGAGAGCAAGGAAGGCA 2078
Db 2217 TCAGTATCTTGGGAAGCTTTAATTTGTGGGATCATCCAAGAAAGAGCAGAGGAAGG 2276
QY 2079 TCCTTTCAGATGATGAAGAGGGGCCAAGGACACTGGGAGGAGACAGTGCACAGAGCAGAG 2138
Db 2277 TCCTTCTTCTGATGAGGAAGGGGACCAAAAGCAATGGAGGAGACCAACAGAAAGCTGAT 2336
QY 2139 GAGGCCAGCAAAAGACAAAGAACCGGAAACAGACGTGTTCTTCTGCCAGACCCAGGAGCAG 2198
Db 2337 GAGGCCGGAAGAAAGACAAAGAGACGGGGAACAGCGGATCCTTGTCTGTTTCCCAAGAACAT 2396
QY 2199 GACCAAGCCAGGAAGTTCTTCCCGAGGAGCGGAGAGGAGGAGGAGGAGGAGGAGG 2258
Db 2397 GATCCAGGGAAGGAGGAGTTCTTCCCGAGGAGAGGAGGAGGAGGAGGAGGAGGAGG 2456
QY 2259 GTCTCCACTTGGAGTCAATTTAAAGATAGTCACTTCCAAAGAAAAAATCAAGTCAAAA 2318
Db 2457 GTTTCACCTGGAGTCAATTTAAAGTTAGTCAAGGAGGAGGAGGAGGAGGAGGAGG 2516
QY 2319 CTGGAAAGAAAGCCGAAGAC-----TCTAGTGTAGAGCAGTTGTCTCACTGAG 2366

Db 2517 CTGGAAGAAAAAGCAAGACTCCATAGCTGCTGTGTATGAAACATTCCACTCCAGAC 2576
Qy 2367 ATCCAGCCGATAGAGAAATCTTGGGTTTCCATTAAAGAAATTCATCCCGAGCGCG 2426
Db 2577 ACTGAAACCCGGTAAAGAAATCTCGGCTCTCAATCAAGAAATTAATCTGTGAGAAAG 2636
Qy 2427 AAGAAAAAGGCAGACGGGAACCAAGAACAGACTGTGGAAGACTCAGGGGCGATGAG 2486
Db 2637 AAGAAAAAGCCAGATGGGAAAACAAGAACAGCCCTGTTGAAGACGAGGGCCACAGGG 2696
Qy 2487 ATTAATGAGAGCAAGCCCAATATGTCCACCGCTGTGCTCTGTCTGATATATATGCACTG 2546
Db 2687 GCCAACGAAATGACTCTGATATCTCCGGGCTGTGCTCTCTGTCTGATATATGATCTGTA 2756
Qy 2547 GAGAGGAGAAAGATGG-----AAGCCCAAGGGAAATACGAGCTGCCCCAGCTGCTG 2597
Db 2757 GAAAGGAGAAATATGAGGACAGCAAGCCCAAAAGGCGCAGACAGCCCAAGCAAGAG 2816
Qy 2598 GGGGCTGTGTACTGTCCGAGAGAGCTCAGTAAGACTCTGTGCTCCACACTGTGAGTCCGA 2657
Db 2817 GCAGCACTGATGTCTCCAGAGAGCTCAGCGAGCTCAGGTTCATATGATGCGACAGCT 2876
Qy 2658 GTCATATGAGGACCAAGGCGAGTCAACAAGTGCAGAGAGCGGTCTCCTTCGTGATATCC 2717
Db 2877 GTGCTGACGGGACGAGGGGCACTTACATTATTAAGAAAGGTCTCTCTTCTTGATATCT 2936
Qy 2718 GCTTCCGTACAGAACTCTTGAACACACAGCGGAGAAAGCCATGSCACCTGTTGAAG 2777
Db 2937 GCTTACGTGACAGAACTCTTGAACAGATGAAGAAAGCGCAGCTGTTAATCAGAGAG 2996
Qy 2778 GTCACTGAAAAAGACATCATTCG---AAGAAACTCTGTGTCTCAACCAAGCTTACCA 2834
Db 2997 GTATTGAAAGAAAGATATGTCAGAAAGAAACCCCAAGGTATCTGAACCTCTGCCA 3056
Qy 2835 GAGGTTAAAGATGGCCATGACAGACATGTGACACAGTGAAGTGAATTTCACTCAGAACT 2894
Db 3057 GAGAAACAGAGAGCCCGGGGCAACAGGTCTGTAAGTGAAGCGGAATTGACCCCGAAGCT 3116
Qy 2895 GTGACAGCCACAGAGACTCAGAGGCTCTCCGTACTGAAGAAATTACCGAAGCATCGGG 2954
Db 3117 GTGACAGCTGCAAGAACTGCAAGGCGCATTTGGGTCCGAAGAAAGAACCGAAGCATCTGCT 3176
Qy 2955 GCCGAAGAGACCAACAGCATGTGTCTCGAGATTTCCCACTGACTGACCTCCCAAGACC 3014
Db 3177 GCTGAAGAGACCAAGAAATGTGTCAAGCAATCTCCCAATTAACCGACTCCCAAGCAC 3236
Qy 3015 ACAGAGGAAGCCACCCAGTTAGAGAGTGAAGAGTGTGTCTAGATCAGAAAGAGAG 3074
Db 3237 ACAGAGAGGCGCACTCCGTGCAAGAGTGAAGAGTGTGCTACCTGACATATGAAGAGCA 3296
Qy 3075 GAGCCCAAGACGCAAGCCATCTCTCAAGCCGTTGACAGACAGGTGAAGAGAGTCCAG 3134
Db 3297 GAGAGGCGGACTCAGAGAGGTCTCTCAAGGCAAGTGGCAAGAAAGTGAAGAGATCCAG 3356
Qy 3135 GTGCTGTCAAC--CAGACTGTGACAGACGCGGTCAAAAGCACTGCAAGAAAGTTGAG 3191
Db 3357 CTGCTGTGCAACCGTGTGGGCGCAGAAAGATGTCTTCAAGCTGTGCAAGACAGAGGCA 3416
Qy 3192 GAGGTGAGAGAGACTCCCAAGTGTGCTTGGAGAAAGAGAGAGAGAGAGTGAAGAGAA 3251
Db 3417 AGACCAAGAGACAGGCTGAAGGCTGAGGCTTGAAGAAAGAGAGAGAGATGATGTTGAAA 3476
Qy 3252 GGAACCGTGCAGAGAGCTGAGAGCTTGTGACAGAGGCTCTGAGACTGAGACAGGCT 3311
Db 3477 GTAGATGCTCAGAGAGCAAAAACTGAGCTTTTACCAAGAGAGAGTGTGTGGGCGAGAC 3536
Qy 3312 ACTCCAGAGAGCCCTTG--AAATTCTGAAGTCAAGCGAGATGTAGC-----CAT 3359
Db 3537 ACCCAAGAAAGCTTTGAAAAAGCTCTCAAGTCAAGAGAGCATAGAGTCAAGTGAAGCTT 3596
Qy 3360 GTCCGCAACTGCA-----GGTATCAGGCTCAGAGAGCTGATGAGAA 3401
Db 3597 GTAACCACTTGTCAAGCCAAACCTTAGCTGGGTAAAAATCAGAGAGATGTGATGAAA 3656

Qy 3402 CAGGCGGTGGCCCTGTGATCCGAAACCTTGACAGACAGTGAACCAATGGAAGCACT 3461
Db 3557 CAGGCTATCCCTCTGCTGTGTGAAGAACCTTACACAGCTGATGATGTGAAGACACC 3716
Qy 3462 CCCTTACAGATTCAGACACTGCAATGGAACAAGCAAGATGAACCAATTGACAGCCAG 3521
Db 3717 CCGGTACCGACTTGTGACGACAGGCAACACCAGAAAGACGAGATGTGGAATCTCAT 3776
Qy 3522 GACAGTAAAGCCACTGCACTGTGTGAGGCACTCAGGTCAACAGAAAGAGGCGGCTACT 3581
Db 3777 GAGGAGATGATGTGTGATCTGTGATCCAGTCAAGGGGCAAGAAAGCAGAGGCACTTCT 3836
Qy 3582 GCTGAGAAAGAGAGCCCTTGACACTACCTAATTAATGTTCCAGCCCAAGAAACATGGG 3641
Db 3837 GCACAGAAAGAGAGGCTCCAGACCTTCCAGTTTGTGTTCCAGAGAAACCTAAGAA 3896
Qy 3642 GAAGAACAGAGAGAGATGTTCTTGAACCTACACAGCAAGAGCTTACTGTGCAAGCCGTG 3701
Db 3897 CAATCAAGATGGAAGACACTTAGAGCATACAGTAAAGAGGTGTCAGTGGAACCTGTA 3956
Qy 3702 CCGGTTCTGCAAAAGACTGAGTGGGTCAAGAGGTGA-----GGTTGACTGTTG 3752
Db 3957 TCCATTTCTGTCAAAAGACTGAGGGGACTCAAGAGGCTGACAGATGCTGATGAGAAAC 4016
Qy 3753 GATGAGAAAAAGCTCAAAAGAAACAGAGAGTGTGTAACCTGG-----ACCAAC 3806
Db 4017 AAAGACGTACATTTTCTGAAGGACTTGAAGGAGTGTATAGCACAGCAATPAACGTACGT 4076
Qy 3807 AGTCAAAAGGCTGTGATGTGACATATGACAGTGAATGAGTGGAGTGGCGGCTGACAG 3866
Db 4077 CGGAAAAAGTCACTGAAGTTGGCCCTTAAAGTGAAGGACAGAAAGCTGAATGTAA 4136
Qy 3867 GAAAAAGAGACTGATGAATGACAG-----TCTTAAGCTGTGAGAGGGA 3911
Db 4137 AAGATGATGTCTTTGAACCTGACAGTCAACGCTAAGCTCTCCATCCCGCTGAGAGGA 4196
Qy 3912 GAGATGAGAACTGACGTTGAAAGAGAAAGGAGCAAAAGCCAGAGCAAGTGAAGAA 3971
Db 4197 GAGATGATGTTCAAGTGTGAAGGAGAGAAACAGAAAGAGCCCAATGTAAATGAA 4256
Qy 3972 GAAGG---TGAGCAGAGAAACAGCCGCTCTGAGCATGAAGAACCTTACGGAAGCCAGT 4028
Db 4257 GAGAAAGCTTGAAGCAAGAAACAGCTGTATCTGAAAGAGTCAAGTGAACAGCTCTC 4316
Qy 4029 CTGACACTTGAACATGCCAGCTCAGAGAGGGGAAAGCATCTGGGAAGCTTTGAGAGAAC 4088
Db 4317 CAGACAGTGAATGTGCCATCATAGATGAGGCAAGAGAAATGACAGAGTTTGAAGAGAAC 4376
Qy 4089 CCTTCTC---TCCAGACCAAGCAAAAGCAAGTGTGACATGAAGTTCAAAGCTG 4145
Db 4377 CCTCTCTCTGCTTGAAGTCAAGAGAGGCAAGTATGCAAAATTCAGTTCAAGGCTCT 4436
Qy 4146 GACACACAGTCACTCAACAGCAGAAAGCTGTGAAAA-----GGTCAATGAAGAGGTT 4199
Db 4437 GAGCATATTTCACTTAAACAGCGCTGCAAGAGAGAGAGAGAGAGTCTTGAAGAACTGCC 4496
Qy 4200 GTGATTTCAAGACAGGTGAAGTCAAGATGTGTAGGTGACACTTAATTAACAGCTGAG 4259
Db 4497 AACATTTTGAAGACAGGTGAAGCTGTGAGCTGCAAGGTGACATTTAGTTCTGGAAGAG 4556
Qy 4260 AAGTCTCTGCAACGAGTGGCACTGCACTTTACAGCATGCAAGAGACACGATACCTCTG 4319
Db 4557 AAATCTCTGAAAAAAGAAAGCTTTCCTCCCTCATCAAGGAGAAAGAGCTGTGTCCCA 4616
Qy 4320 GGGCCTGAGCTCAGGCGAGAAATCCATCCATCAATGATTAATCTGCTCTGAAAGCAC 4379
Db 4617 GGGCCGCACTGTCAAGCAAAATTCACACAGTATGATGTGTCTTCACTCAAGAAAGAGC 4676
Qy 4380 CTACATCTGCACTTCAAGAGAAATTAAGCGCATCCAGAGAGAGCGATCAGAGAGAG 4439
Db 4677 TTAAGTTCGACCTGGAAGGAGAGAAAAACCATCATCAGAAAGTGAAGTCAAGTGAAGTC 4736

Hortigan, S., Soppet, D. R. and Weaver, Z.
Cancer gene determination and therapeutic screening using signature
gene sets
Patent: WO 0194629-A 275 13-DEC-2001;
Avalon Pharmaceuticals (US)
Location/Qualifiers
1. 6608
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 2091 a 1418 c 1827 g 1272 t

Query Match 34.8%; Score 2144.4; DB 6; Length 6608;
Best Local Similarity 64.3%; Pred. No. 0; Mismatches 1796; Indels 543; Gaps 45;
Matches 4215; Conservative 0;

3 GGGCCGCGGAGTAGAAGACCACTGAGCCATGGCGGAGGAGGAGTTCACCCGAGGAGCGG 62
162 GGCTAGGCGCGGAGAAATCGCGAGAGCCATGGCGCGCGGAGGCTCCACCGAGGAGCGC 221

63 AGCCCGGAGAGCCG--GCGGGAGGAGCACGCGCGAGCGAGTGTCTCATGTGCCAT 119
222 AGCCCGGAGAGCCGCGCGAGGAGCTCCAGCGCGGCTGAGCCGAGCCGAGCGCGAGC 281

120 GGGCCGCGAGCTGAAGCTCTGGGA-----GCAGCTGGAGACCCGCGGAGCGGAGC 170
282 GGGCCCTGGCGGAGGCGGCGGAGACCAACCGCGGAGCCCGCATGCTGCTCGGAGC 341

171 CCGCGCACCAAGCTCCCAAGAAATGGCGAGCTGTCTTGTCTGACGCGCTAGCTGAA 230
342 CCGCGCACCAAGCTCTCTACAGAAATGGCTGCTGTCCACATCAATGGCTTACTAG 401

231 CAAGGAGATGTCATGTCCAGAGAGAAACAGAGG----- 268
402 CAAGATGAGCTCAGCTCCAGAGGAGTGAACCTAAATGGCCAGAAAGAGAGCTGAGCGGT 461

269 -----GCAGAGGAGAAAGTCTTGATGAGAGTGTGAGACAGCGA 308
462 CAAGAGCCCTAAACAGCCAGGAGGAGAAAGATGTCATGTCACGAGAGTTGAGACAGGA 521

309 GAGTCAGAAATGTGAGAGAAACCGAGTTGAAAGAAATGGCGCCCACTCCACAGCT 368
522 GACTCTGAAGATGTGAGCAAGAAAGACTCCGATTAAGAGATGGCTACTAAGTCAAGCGGT 581

369 GTTGAAGATATCAAAAGAGTGGGAGAGAGACATCAAGAAATTAATGAACAATCCCT 428
582 GTTCAAGACATCAAGATGATGGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 638

429 GCTTCAGAAATATGTGAGAAATGTGAACGCTGCTGAGTCCGAGCTAATGATGT 488
639 TCTTCAGAAAGATTTAAGAGCTAACCAACCACTGAGTCCGAGCTAATGATGT 698

489 GGCTCAAGAAAGTATTTAAATTTGTGTTTAAATTTACGCGTGAAGAGATTAATAAT 548
699 GATTTAAGAGAGTGTTTAAGTTTGTGGCTTTAAATTTACGTGAAGAAAGATTAAGAA 758

549 GAAAGTCAAGATCTGTCACTACTCTCAAGAGAGATGAAGCGAGGCGAGAA 608
759 GAGAAAGCTGACACTGTCCAGTCACTGTGAAGAAAGATGAAGGAGGAGAGAGAGCA 818

609 GCCTCTGTGAGCTGAGAGACCAAGAGCCGAGTGTGAGAGCTGCCGTCCGAGAGTGA 668
819 -----GGGCGTGGCCACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 863

669 GCATCCAAAGAAAGTGAAGAGATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 728
864 GATCCAAAGAAAG 923

729 GAACAG 788
924 GAGCAAG 980

789 GCCAAAGATGAAG 848

849 AGCAGCCAGTCAACAGTGAAGACATCTTCTTCAAGAAATTTCTTCACTCACGGTTGG 908
1041 ACTAGTCCCGTGACAGTGAAGACAGATCAACCTTCAAAAAATTTCTTCACTCAAGTTGG 1100

909 GCGGCTGGCGCAAGAGACCACTTCAAGAAATCAAGAAAGATGATCTGGAAGCTGCC 968
1101 GCGGCTGGCGCAAGAGACCACTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1160

969 GAGAAAG 1028
1161 GAGAAAG 1220

1029 CCAGCTCGAGAGAG-----GCAGAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAG 1061
1221 GTTGCCTCGAG 1280

1062 ACAG 1121
1281 GCGCAG 1340

1122 GTTGTGAGCTGAG 1181
1341 GTCACTGGCTCGAG 1400

1182 GATGAAGAGATGAAG 1238
1401 GATGAAG 1460

1239 GAGAAAG 1298
1461 GAG 1496

1299 GGAACAG 1358
1497 GAAACAG 1556

1359 GCTAGAGCTGAG 1418
1557 GCGCAAGCTGAG 1616

1419 ACTCACTGAAG 1478
1617 ACACAG 1676

1479 GTCACTGAG 1538
1677 GTGAGTGAAG 1736

1539 AAGAACTCTTCAAG 1598
1737 AAGAACTCTTCAAG 1796

1599 CGAG 1658
1797 AG 1853

1659 GAGAGTCTGATGAAG 1718
1854 GAGAGTCTGATGAAG 1913

1719 ACCAGTGTCTGAAG 1778
1914 ATCAAGTGTCTGAAG 1973

1779 ACTACTTCTGATGAAG 1838
1974 GCTACTTCTGATGAAG 2033

1839 ATGTGACACCAAG 1898

Db 2034 ATGGTGACGCCAAGAACGGTGTAGACGGCTTTCGGAAAGTGATAAAGAGATGAGCTG 2093
Qy 1899 GAGAAGGTCAAGAGCGCCACCTTGTCTCTCACTGATAGCAAGTGTGAGAAATCAAGAT 1958
Db 2094 GACAAGGTCAAGAGCGGTACCTTGTCTTCCACCGAGAGCACAGCTCTTGAATGCAAGAA 2153
Qy 1959 GNAGTCAAAACTGTTGGTAGGNAACAAGCCAGAGNACCAAGACCTAGGTTGGATACT 2018
Db 2154 GAAATGAAGAGGCGGTGGAAAGACCAAAAGCCGGAAGAACCAAGCGCAAGGTGGATACC 2213
Qy 2019 TCAGTGTCTTTGGGAAGCACTGATTTGTGTGGATCATCTCAAGAAAGAGAGCAAGGAAGCA 2078
Db 2214 TCAGTATCTTTGGGAAGCTTTAATTTGTGTGGATCATCTCAAGAAAGAGCAAGGAAGG 2273
Qy 2079 TCCTCTTCAGATGATGAAGAGGCGCAAGGACACTGGAGGGGACAGTCAAGAGCAGAG 2138
Db 2274 TCCTCTTCATGATGAGGAGGGGGACCAAAAGCAATGGAGGAGACCAACCAAGAAAGCTGAT 2333
Qy 2139 GAGGCCAGCAAGAACAAAGAGCCGGACAGACGCTGTTCTGTCAGCACCCAGAGCAG 2198
Db 2334 GAGCCCGAAAGAACAAAGAGACGGGACAGACGGGATCTTGTGTTCCCAAGAACAT 2393
Qy 2199 GACCAAGCGCAAGGAAGTTCTCAACCGAGCCAGCGGAAGCCCTTCGGAAGGGGAAGGT 2258
Db 2394 GATCCAGGCGAGGAAGTTCTCTCCCGAGCAAGCTGGAGCCCTACGGAAGGGGAGGC 2453
Qy 2259 GTCTCCACTTTGGGAGTCAATTTAAAGATTAGTCACTCCAGAAAAAATCAAGTCAAAA 2318
Db 2454 GTTTCCACTTTGGGAGTCAATTTAAAGGTTAGTCAAGCCCAAGAAAAAATCAAGTCCAAG 2513
Qy 2319 CTGGAAGAGAAAGCCGAGAC-----TCTAGTGTAGACGATGTTCTCACTGAG 2366
Db 2514 CTGGAAGAGAAAGCCGAGACCTCATAGCTGGGTCTGGTGTAGAACATTCACCTCCAGAC 2573
Qy 2367 ATCGAAACCGATAGAGAAGATCTTGGTTTCCATTAAAGAAATTCATCCCGGACGCGG 2426
Db 2574 ACTGAAACCGGTAAAGAGAAATCTGGGTCTCAATCAGAAGATTATTCCTGGAGAGG 2633
Qy 2427 AAGAAAAGGCGAGCGGGAAGCAAGAACCAAGCCACTGTGGAAGACTCAGGSCCAGTGGAG 2486
Db 2634 AAGAAAAGGCGAGATGGAAACAAAGAACCAAGCCCTGTGGAAGCGCAGGSCCAACAGGG 2693
Qy 2487 ATAAATGAGNAGACCTTAATGTCTCCAGCGTCTGTGCTCTGTCTGATGATATGCAAGTG 2546
Db 2694 GCCAACGAAGATGACTGTGATGTCCTCGGCGGTGTCTGTCTGTGATGATGCTGTA 2753
Qy 2547 GAGAGGAGAGAGATGG-----AAGCCAGGGGAATACGGAGCTGCCACGCTCTG 2597
Db 2754 GAAAGGAGAAATGGAGGCAACAGCCAGCCAAAAGGCGCAGAGCCCGCAGAGAAG 2813
Qy 2598 GGGCTGTGTACGTGTCCGAGGAGCTCAGTAAGACTCTGGTCCACACTGTGAGTGTGCCA 2657
Db 2814 GCAGCCACTGAGGTGTCCAAGGAGCTCAGCGAGAGTCAAGTTTATATGATGGCAGAGCT 2873
Qy 2658 GTCAATTGATGGACAGGCGAGTCAAGAGTGTGCAAGAGCGGTCTCTTGTGATATCC 2717
Db 2874 GTCTGTGAGGAGCGAGGCGAGTACCATTTATGAAAGAGGTCTCTCTTCTTGGATATCT 2933
Qy 2718 GCTTCCTGTAAACAGAACTCTTGAACACACAGCGGGAGAGCCATGCCACCTGTTGAAAG 2777
Db 2934 GCTTCAGTGACAGAACTCTTGAACAAAGTGAAGCTGAAGCCGACACTGTTAACTGAGGAG 2993
Qy 2778 GTCACTGAAAAGACATCAATTGC---AGAAGAACTCTCTGTGCTCAGCCAGACCTTACCA 2834
Db 2994 GTAATTGAAAAGAGAAGTAAATTGCAGAGAAGAACCCCCCAAGTTACTGAACCTCTGCCA 3053
Qy 2835 GAGGGTAAAGATGCCATGACGATGTTCAACAGTGAAGTGGATTTCACTCAGAAAGCT 2894
Db 3054 GAGAACAGAGAGGCGCGGGGCGACAGGTCTGTAGTGGCGGAATTGACCCCGAAGCT 3113
Qy 2895 GTGACGCCACAGAGACTCAGAGGCTCTCGTACTGAGNAGTTACCGAAGCATCGGG 2954
Db 3114 GTGACAGCTGCAGAACTGCGAGGGCCATTGGGTTCCGAAGAAGGAACCGAAGCATCTGCT 3173

Qy 2955 GCCGAAGAGACCAAGACATGGTGTCCGCAAGTTTCCAGCTGACTGACTCCCGCAGACACC 3014
Db 3174 GCTGAAGAGACCAAGAAATGGTGTCCAGAGTCTCCAGTTAAACCGACTCCCGCAGACACC 3233
Qy 3015 ACAGAGGAGCCACCCAGTTCCAGGAGGTAGAGTGGTGTCTAGATACAGAGAGAG 3074
Db 3234 ACAGAGGAGCCACTCCGGTGCAGGAGGTGGAAGGTGGCGTACTTGACATAGAGAGCAA 3293
Qy 3075 GAGCGCCAGACGCGAGGCCATCTCTCCAGCGTTTGCAGACAAGGTGAAAGAGGATCCAG 3134
Db 3294 GAGAGCGGACTCAAGAGGTCTCTCCAGGCAAGTGCAGAAAAGTGAAGAGGAATCCAG 3353
Qy 3135 GTGCTGTCAACC---CAGACTGTGCAGAGAAACGGGGTCAAAAGCACTGGAGAGAGTTGAG 3191
Db 3354 CTGCTGTGCCACCGGTGGCCAGAAAGATGTCTTTCAGCTGTGCAGAGCAGAGGCGAGAA 3413
Qy 3192 GAGGTAGAGGAGGACTCCGAAGTGTGGCTTCGGAGAAAGAGAGGAGTATGCCGAAA 3251
Db 3414 AGACCAAGAGAGGCTGAAGCGTGGGGTCTGAAGAAAGAGACGGATGTAGTGTGAAA 3473
Qy 3252 GGACCCGTGCAGGAAGCTGGAGCTGAGCATCTTTCAGCAGGGCTCTGAGACTGGACAGGCT 3311
Db 3474 GTAGATGCTCAGGAGGCAAAAATCTGAGCTTTTACACAGGGAAGTGGTGGGCAACACC 3533
Qy 3312 ACTCCAGAGAGCCTTG---AAGTTCTGAAAGTCAOCCGACAGATGTAGAC-----CAT 3359
Db 3534 ACCCAAGAAAGCTTTGAAAAGGCTCTCAAGTCAAGAGAGCATAGAGTCCAGTGAAGCTT 3593
Qy 3360 GTGCCACGTCGCA-----GGTTATCAAGCTCCAGCAGCTGATGGAA 3401
Db 3594 GTAACTACTTGTCAAGCCGAAAACCTTAGCTGGGGTAAAATCAAGAGAGATGGTATGGAA 3653
Qy 3402 CAGCCGTGCCCCCTGAGTCAATCCGAAACCTTCAGACACAGTGCAGACAAATGGAAGCACT 3461
Db 3654 CAGCTATCCCTCTGACTCGGTGGAAACCCCTTACAGACAGTGAAGTGTGAGAGCACC 3713
Qy 3462 CCCTTAGCAGATTCAGACACTGCAGATGGGACACAGCAAGATGAAACCATTTGACAGCCAG 3521
Db 3714 CCGTAGCCGACTTTCAGCGCACAGGCGCAACCCAGAAACAGAGATTGTGGAAATCCAT 3773
Qy 3522 GACAGTAAGCCACTGAGCTGTCCGAGTCAAGTGCAGAGTGCAGAGAGAGGCGGCTACT 3581
Db 3774 GAGGAGAAATGAGTGCATCTGTGATCCCACTCAGTCCAGGGGCAAGAGCAGAGGAGTTCCT 3833
Qy 3582 GCTCAGAAAGAGGAGCTTCGACACTACCTAAATGTTCCAGCCCAAGGAAGAACATGGG 3641
Db 3834 GCACAGAAAGAGGCTTCAGCACCTTCCAGTTTGTGTTCCAGGAAGAACTAAAGAA 3893
Qy 3642 GAAGAACCCAGGAAGAGATGTTCTTGAACCTTACAGCAAGAGCTTACTGCTGCAGCCGTG 3701
Db 3894 CAATCAAGATGGAGACACTCTAGAGCATACAGATAAGAGGTGTCAAGTGAAGAACTGTA 3953
Qy 3702 CCGTTCTGCAAGACTGAGTGGGTCAAGAGGGTGA-----GGTTGACTGGTTG 3752
Db 3954 TCCATTTCTGTCAGAAAGCTCAGGGGACTCAAGAGGCTCAGAGTATGCTGATGAGAAAC 4013
Qy 3753 GATGGAGAAAAGTCAAAAGAAAGAGGAGGTGTTGTACACTCTGG-----ACCCAAC 3806
Db 4014 AAAGACGTACCATTTTTCGAAGGACTTTGAGGGTCTATAGACACAGGCAATAACAGTCACT 4073
Qy 3807 AGTCAAAAGGCTCTGATGTGACATATGACAGTGAAGTGTGAGGTGGCGGGTGTTCAG 3866
Db 4074 CGGGAAGAGGTCACTGAGTGGCTTAAAGGTGAAGGACAGAGAGTGAATGTAAA 4133
Qy 3867 GAAAAGAGAGTACTGAAGTGCAGAG-----TCTTAGCTGGAGGAGGGA 3911
Db 4134 AAGGATGATCTCTTGAACCTGCAGAGTCAAGTCAAGTCTCTCCATCCCCCGTGGAGAGA 4193
Qy 3912 GAGATCGAACTCACTTGAAGAGGAGAAAGGAGACAAAGCCAGAGCAAGTGAAGTAA 3971
Db 4194 GAGATGATGTTCAAGTCGAAGAGGAGAAACAGAGCAGAGCCCAACCCATGTGTAATGAA 4253

QY 3972 GAAG---TGAGCAGAAACAGCGCTCTTGAGCATGAAGCACTACGGAGGCCACTC 4028
 DB 4254 GAGAGCTTGAGACAGAAACAGCTGTACCTGATCTGAAGAGTGAAGTAAGAGCTCTC 4313
 QY 4029 CTGACATCTGCAATGCCAGCTCAGAGAGGGAGCACTGGAGCCTTGGAGAGC 4088
 DB 4314 CAGACAGTGAATGTGCCATCATAGATGGGCAAGAGAGTCAAGATTTGGAGAGAGC 4373
 QY 4089 CTTCTC---TCCAGACCAAGACAGAGTTGATAGAGTTCAAGTTCAAGCCTG 4145
 DB 4374 CCTCTCCCTGCTAGGTCAGAGAGGACATGATGCAACAAATTCAGATTCAAGCTCT 4433
 QY 4146 GACACAAAGTCACTCAACACAGCAAGCTGTGAAA-----GTCATGAAAACGTT 4199
 DB 4434 GAGGATCATTTCACTCTAAACAGCGCTCAGAGAGAGAAAGTTTGGAGAAACCTCC 4493
 QY 4200 GTGATTTGAGAGACAGAGAAAGTCCAGAGTGTAGTGACACATTATACCACTGAG 4259
 DB 4494 AACATTTTGAAGAACAGGTGAAACGTTGAGCCTGCAAGTGACATTAGTTCTGAAAG 4553
 QY 4260 AAGTCTCTGCAACGAGTGGCCACTGACTTTTACATGACAGAGACACGGTACCCTG 4319
 DB 4554 AATCTCTGAAAAAATGAGACTTTTGCCCTCATCCAGGGGAGATGCTGTGCCACA 4613
 QY 4320 GGGCTGAGTCTCAGGAGCAATCCATCCCATATATATATCTCTGCTTGAAGACAC 4379
 DB 4614 GGGCCGCACTGTCAGGCAAAATGACACAGATATATCTGCTATCAACAAAGAGGC 4673
 QY 4380 CTACATCTGACCTCAAGAGAAATTAAGCGCATCCAGAGAGACGATCAGAGAGAG 4439
 DB 4674 TTTAGTTTCGACTTGAGAGAGAGAAACCATATCATCTGAAGTGAAGTCAATATATC 4733
 QY 4440 GACAAAGCCAGATGCTGTCTGATGCTGACGCAAGAGAGATGACCAATCGAAAAAGTC 4499
 DB 4734 GATAGGAGGTTGCTTGCC-----AGAGGTCAAAGTAGATGTGCAATTTG---AGAT 4784
 QY 4500 CTCAAGCTGAACCTGAAGATCTGGAATTGAGAGTAAGAGCAACAATTTGTGTGAC 4559
 DB 4785 TTAGAGCTCGAAATGGGATTTTGGAACTTGAGACCAAAAGCAGTAACTTTGCCAAAAC 4844
 QY 4560 GTCATTCGACAGCGGTTGACCACTGCGACGTAC---AGAAACAGCCCCGAAACTCAT 4616
 DB 4845 ATCATTCGACAGCGGTTGACCACTGCGACGTACCAAGAAACAGCCGAAATTTTG 4904
 QY 4617 GCTTATGATTCAAGACCCAGGTTT-----CTGATGACGCTTGAACG 4660
 DB 4905 ACGTGTAGTTACAGACCAAGCTCACGTGATAAAGCTGACAGCAGGACGCTGACAG 4964
 QY 4661 CAGGAGCCCAACAGATGCTGACAAAAATGAAGATGCAAGATGAAACACCCAGTGC 4720
 DB 4965 GAAACGAGAAAGAGAGAGAACTCAGGCTCTGCAAGAGATGAAACACCAATTAAT 5024
 QY 4721 GAGGCCA-----GAGAGACTTGGAAAGTCCGACGCTGAG 4759
 DB 5025 TCAGCAAAAGAGAGATCAGAGTCAACCGAGTGGACAAAGCAATTTGATATTTTCCAA 5084
 QY 4760 GGCATAGGCTCAGCCTCGGAAATGCTTGGCGGCTTGGAGTTGAAGCGCCGAGTCA 4818
 DB 5085 GACATGAGTGAAGCTCAGAAAAAGACATGACTGTTGAGGTGAAGGTTCCACTGAAAT 5144
 QY 4819 ----- 4818
 DB 5145 GATCAGAGCTGGAAGAGGTCTCTCCCATCTGAGAAAGAGAGAGTGAAGCTGGAACA 5204
 QY 4819 -----AGTAAGCATTTGAAAGCTGCT 4840
 DB 5205 AAGTGTGTCAGAAAGATGATGCTATGCTTTGTAAGCAAAAGATAGAGAGATCATCA 5264
 QY 4841 CCTCAACCCAAAGA--TCCAAAGAGAGATGCTGCTGATGAGCCTCAGCTCCAAAGCTTA 4898
 DB 5265 GTTGAACCGAAAGAGATGAAAGATGATGATGTTGATGATGATGATGATGATGATGATG 5324
 QY 4899 GCCCAGGAGAGGCGAGTCTCTGGAACCTTAACCAAAAGAAATCCCCAGACACCAAGCA 4958

DB 5325 GCTTGGCTGATATGATGCTCTGAGAGGCTTTAACCAAAAGTCTCCAGATTAACAATGGA 5384
 QY 4959 CCAAGCTTAACCGAGAGGCGATCCCCCAAAAGTTGAGTCCAGAGAAAGAAATGCT 5018
 DB 5385 CCAAAACAAAAGAGAGAGATGCCAGAGAGTAAATTTGACGAAAGGAAAGTGCAC 5444
 QY 5019 ACCAAGTCAAGTCAAAAGAAACAGAGCCAGAGAGAGAGAGATCTGAGAGCCAAAGGA 5078
 DB 5445 AGTAATCAGATTAAGGATCAACCCCAAGCAAGAGAGATTAAGAAACAAAGAGA 5504
 QY 5079 GACCTGGAGAAATCTAAGATGTTAGTTG-----CTCATTTGAC 5117
 DB 5505 GAATCTGAAAGTGAACCTTAAGAAATCTTAAACATCATGACATTAACTCATTTGCT 5564
 QY 5118 ATCTGAAGACAGAAATGTAAGAAACAGTCAAGAACAA-----GATGCTGTGTGGAC 5173
 DB 5565 GTTTGAAGACCAAGATGTGAAGCAAGTATGAAAGAAATGATGCTGCTGAGAG- 5623
 QY 5174 CTTGAGACCAAGATTTTCAAGCCCATGAGATCCAGAGAGAGAGGCGCTCCAAATGATTC 5233
 DB 5624 CTGAAGACCAAGTATTTTCAAG-ACTTGAGAAATGGAAGAGAGGACATCAACTGATCTCA 5682
 QY 5234 ACCAGTAGAGACACCCGCAATTTCTGAGGCTCATGCGGAGCTAGAGCCAGTAACTT 5293
 DB 5683 TTTCTAGAGAC-CCCTGACAACTCTGAGCTTATGAGAGCTAGAGCCATTTAACTTT 5741
 QY 5294 TCTCTGTTCAAGACTGCTTGAATTTGCCCTTGAATGCGGTCCGTATTTCTAACAT 5353
 DB 5742 TCTCTTTTCCAAAGCAACCTACAAATTTTCCCTTGATAC----- 5781
 QY 5354 TCTCTGTTCAAGACTGCTTGAATTTGCCCTTGAATGCGGTCCGTATTTGCGATTTA 5413
 DB 5782 -----CATTAATATTCGATTTA 5799
 QY 5414 AGGTCTGCGCTTCAACTGGAAC-CAATTCGCACTACTAGTCACTTCAACT 5472
 DB 5800 AGGTCTTAATTTCTTAACCTGGAACGAGTGGCAATCTTAATTTCTGCTTGAACCT 5859
 QY 5473 GAGACATCTCTCTTAATGATTTATATGATGATGATGATGATGATGATGATGATGAT 5532
 DB 5860 GAGATATCATTTCTTAATATTTATATGATGATGATGATGATGATGATGATGATGATGAT 5916
 QY 5533 TGTATATTTTCTTAACGTT-----AAGCATGCTTTTGTATTA 5575
 DB 5917 TGTATATTTTCTTAATGTTTAAAGAAATGTCAGAGATCTCATGAGCTTTTGTATCA 5976
 QY 5576 TGCATATATTAAGGAGTGTGACCATAGAGAGCTTTGAAAGCTCCAGAGCTCAACTG 5635
 DB 5977 CACAGTATATGATGGGAGATGTCATAGTGCAGGCTTGGAGCTTTAAAGCTCAAGTTA 6036
 QY 5636 TAACCTGAGCAACAGATTAACATTC-----CTGGCAAGAGAGACAGATCTTTT 5686
 DB 6037 TATAACCAACAAAACAGAGCTCTTAAGATGTAACATTCCTGATCAAGATCAATTCCT 6096
 QY 5687 TTAAGTTTACTGATGCTTGAATCTGTGGGCTTGAATCTCTGAAAGTGTGTTTCTCT 5746
 DB 6097 TAAATTCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6156
 QY 5747 ATGCACAGCAGCTCAGAAATTAAGAAACCCCAATTTGAAACATCCAGAGATGCCAATTT 5806
 DB 6157 ATTAACAGGAGTGTGCCAAACCTAAAGAGATTTTGAACATTAAGAGATGTTATTTGTC 6216
 QY 5807 AACATGATTTTTCCTCCCTTTTGTCTAATCCAGTCCAGGTTGAAAGAGATCTCTCTG 5866
 DB 6217 ATGGGAAATTTTG-----TTTCAACCAAGTGGAGGTTAAGAAAGATTAATTTCT 6269
 QY 5867 TGTCAAG-----TTAAGCCTGTCTTATATGATATGACAAATGAGT 5909
 DB 6270 GGTAGCAATTAATTAATCATCTTTTCTACTGTTATGTTGTTGAGCCAGATAGT 6329
 QY 5910 GTGCTTAAGCGCATGAG-----TGTTCCTAATGCAAGAGAAATCTGT 5953

Oy	1419	ACTCACTGACACACCTAAAGTCTCTGAAAGAAAGACGCTGGCCAAACACCCAGAAAGGCATT	1478
Db	1617	ACACAGGAGACTGACCTCACTCTCTGATGAAAGGTGCTGTCCAAACCCCCGAAAGGCGTT	1676
Oy	1479	GTCAGTAGAGGTGAGATGCTGTCTCTCAAGAAAGAAATCAAGTACAGGAAATCCCTTG	1538
Db	1677	GTCAGTAGAGGTGAAATGCTGTCTATCAAGAGAGAAATGAAGTGCAGAGGAATGCTCACTA	1736
Oy	1539	AAGAAACTTTCAGTAGCTTAAGCTTTAAAGAGCTGTCTGGGAAGAACACAAAGGGGAAA	1598
Db	1737	AAGAAGCTTTTTCACACACTGGCTTTAAAAAGCTTTCTGGAAAAAACAAGAAAGGGAAA	1796
Oy	1599	CGAGAGAGTGGGGGAGACGAAGAGCTCTGAGAAATACCAACATTCACACGGAATCCCCA	1658
Db	1797	AGAGGA---GGAGAGACGAGAAATCAGGGGAGCACTCAAGTTTCAGCCGATTCCTCCG	1853
Oy	1659	GAGAGTCTGATGACGAGAAAGGAGAGCTGTGCTGCTCCCTCCGAGAGAGCTTGAGAG	1718
Db	1854	GACAGCCAGAGGAGGACAAAGAGGCGAGAGCTGTGCTCTCCCTGAGAGAGCCCGAGAG	1913
Oy	1719	ACCAAGTCTCTGGAGAAAGGGCCGCTGGAAAGCACCCAGAAATGGGGAAAGCTGAGGAAGA	1778
Db	1914	ATCACGTCTCTGAAAAAGGCGCTTAGCCGAGGTGACAGAGATGGGGAAGCTGAAAGAA	1973
Oy	1779	ACTACTTCCGATGAGAGAAAGAAAGAGAGGAGTCACTCCCTGGGCAATCTTTCAAAG	1838
Db	1974	GCTACTTCCGATGAGAGAAAAAAGAAAGATGTCACTCCCTGGGCAATCATTTCAAAG	2033
Oy	1839	ATGGTGACACCCAAAGAAACGGGTCCGAAAGCTTTCTGAAGTGAACAAGAGGAAGAGCTG	1898
Db	2034	ATGTGAGCGCCAAAGACGTGTGTAGACGCGCTTCGGAAAGTGAATMAAAGATAGAGCTG	2093
Oy	1899	GAGAAGGTCAAGACGCGCACCTGTCTCTCACTGATAGACAGTGTCAAGAAATGCAAGAT	1958
Db	2094	GACAAAGTCAAGAGCGCTACTTGTCTTCCACGAGAGCACAGCTCTGAATGCAAGAA	2153
Oy	1959	GAAGTCAAACTGTTGTGAGGAACAAAGCCAGAGAAACCAAAGCTTAGGTTGATGTCT	2018
Db	2154	GAAATGAAAGGAGACGTGGAAGAGCCAAACCCGAAAGAACAAAGGCGCAAGGTGATATCC	2213
Oy	2019	TCAAGTCTTGGGAAAGACGATTTGTGTGGATCATCCAAAGAAAGACCAAGAAAGCA	2078
Db	2214	TCAGATATTGGGAAGCTTTATTTATTTGTGTGGATCATTCAGAAAGAGCAAGAGAAAG	2273
Oy	2079	TCTCTTTCAGATGATGGAAGAGGGGCCAAGGACACTGGAGGGGACAGTCAAGACAGAG	2138
Db	2274	TCTCTTTTGATGAGAAAGGGGGACCAAAAGCAATGGAGAGAGACCAACAGAAACTGAT	2333
Oy	2139	GAGGCCACCAAAAGCAAAAGAACCCGGAACAGACGCTGTTCTGTCCAGACCTCAAGAGCAG	2198
Db	2334	GAGGCCGGAAGAAAGCAAAAGAGACCGGGGACAGAGGGGATCTCTGTGCTTCCAAAGAAAT	2393
Oy	2139	GACCAAGGCAAGAGAGTCTCTCACCCGAGACCAGCGGAAAGCCCTTCCGAAGGGAGAGT	2258
Db	2394	GATCCAGGGCAGAGGAATGCTCTCCCGGAGCAAGCTGTGAAGCCCTTACGAAAGGGAGGGC	2453
Oy	2259	GTCTTCACTTGGGAGTCAATTTAAAGATTGATCACTTCAAGAAAAAATCCAGTCAAAA	2318
Db	2454	GTTTTCACTCGTGGAGTCAATTTAAAGTTTGTGTACGCGCAAGAAAAAATCAAAAGTCCAG	2513
Oy	2319	CTGGAAGGAAGAGCCGAAGAC-----TCTAGGTATAGAGAGTGTTCACCTGAG	2366
Db	2514	CTGGAAGGAAGAGAGATCTCATAGCTGGGTCTGTGTGTAGACATTTCACTCAAGAC	2573
Oy	2367	ATCGAAGCGATAGAGAAAGATCTTTGGGTTTTCATTTAAGAAATTCATCCCGGAGCGCG	2426
Db	2574	ACTGAACCGGTAAAGAAAGATCTCTGGGTCTCATCAAGAAAGTTTATTTCTGTGACGAAG	2633
Oy	2427	AAAGAAAGGCGAGACGGGAGCAAGAACAGACCTGTGGAAAGCTCAGAGGCCAATGAG	2486
Db	2634	AAAGAAAGGCGAGATGGGAAACAAAGAACAGCCCTGTGTGAACGCCAGGGCCACAGGG	2693

QY	2487	ATTATATAGACACACCTTATATGTCCAGCGCGGTCTCTGTGAGTAAATAGCATG	2546
Db	2694	GCCAAACAAGATGATCTGTATGTCTCCGCGCGTGTCTCTGTGTGATGATGTCTGTA	2753
QY	2547	GAGAGGAGAGAAATGTG-----AAGCCAGGGGAAATACGAGCTGCCCATCTCTG	2597
Db	2754	GAATGGAGGAAATATGAGGGACACGACGACCCAAAAGGCGCAGAGCAGCCGACAAAG	2813
QY	2598	GGGGCTGTGATCGTGTCCGAGAGCTCAGTAAAGCTCTGTGTCCACACTGTGATGTCCGA	2657
Db	2814	GACACCACTGAGGTGTCTCAAGGAGCTCACGAGAGCTCAGGTTTCAATATGATGGACGAGCT	2873
QY	2658	GTCATTGATGGGACGAGGGAGTCAACAATGTCCAAAGAGCGGTCCCTGTGGATATACC	2717
Db	2874	GTCCGTACCGGACGAGGGGACGTTACCATTTATTGAAGAAAGGTCTCTCTTGTGATATCT	2933
QY	2718	GCTTCCGTAACAGAACCTCTTGAACAACAAGCGGGAGAGAGCATGCCACTGTGTAAGAG	2777
Db	2934	GCTTACGATGACAGAACCTCTTGAACAAGTAAAGCTGAAGCCGACCTGTTAATCTGABAGAG	2993
QY	2778	GTCACTGAAAAAGACATCATTTGC--AGAAGAACTCTGTGTCTCACCCAGACGTTACCA	2834
Db	2994	GTAATTGAAAGAGAAATTAATTGAGAAAGAAACCCCCACGGTTACTGAACCTCTGCCA	3053
QY	2835	GAGGGTAAAGATGTCCCATGACACATGTGCACCACTGAAGTGAATTTACCTCCAGAAAGCT	2884
Db	3054	GAGAAACAGAGAGGCCCGGGCGCACAGGGTCTGTATGTAGGGCGGAATTAACCCCGAAGCT	3113
QY	2895	GTGACAGCACAGAGACCTCAGAGGGCTCTCCGTACTGAAGAAATTAACGAGACATCGGGG	2954
Db	3114	GTGACAGCTGCAGAAACTGCAGGGGCCAATTGGGTTCCGAAAGAGAACCGAAGCATCTGTCT	3173
QY	2955	GCCGGAAGAACCAACAGACATGTGTCCGACATTTCCCACTGACTGACTTCCCAAGACACC	3014
Db	3174	GCTGAAGAGACCAACAGAAATGTGTGTACAGTCTCCAGATTAAACCGACTCCCAACAGACACC	3233
QY	3015	ACACAGAAAGCCACCCCATTCAGAGAGTAAAGAGTGTGTCTAGATACAGAAAGAGAG	3074
Db	3234	ACAGAGAGGGCCACTCCGGTGCAGAGAGGTGAAGGTGGGTGTAACCTGAATAGAAAGACCA	3293
QY	3075	GAGCGCCAGACGCAAGGCGCATCTCCCAAGCGATTGCAGACCAAGSTAAAGAGAGAGTCCAG	3134
Db	3294	GAGAGGCGGACTCAAGAGGTCTCTCAAGGCAGTGTGCAGAAAAATGAAGAGAAATCCAG	3353
QY	3135	GTGCTTCGCAACC---CAGACTGTGCAGAGAACGGGGTCAAAAGCACTGAGAAAGGTTGAG	3191
Db	3354	CTGCGCTGGCACCGGTGGGCCAGAAAGATGTGTTCAAGCTGTGCAGAGACAGAGGCGAGAA	3413
QY	3192	GAGGTAAAGAGAGGACCTCCGAATGTGTCTGGAGAAAGAGAGAGAGCTTATGTCCGAAA	3251
Db	3414	AGACACAGAAAGCAGAGGTGAAGGCTGTGAAGGTCTGTAAGAAAGAACCGAGATGTATTGAAA	3473
QY	3252	GGACCCGTCGAGGAAGCTGGAAGGTGAGCAATCTTGCAACAGGGCTCTGAAGCTGGAACGGCT	3311
Db	3474	GTAAGTGTCTCAGAGAGGCAAAACCTGAGCCTTTTACAAAGGAGAGGTGTGTGGGCGACCC	3533
QY	3312	ACTCCAGAGACCTTTG---AAGTTCCTGAAGTCAACGCGCAGATGTAGAC-----CAT	3359
Db	3534	ACCCAGAAAGACTTTGAAAAAGCTCTCAAGTCAACGAGATGAGATGAGTCAAGTGAAGCTT	3593
QY	3360	GTGCGCACGTGCCA-----GTTATCAAGCTCCAGACGCTAGTGGA	3401
Db	3594	GTAACCACTTGTCAAGCGCAAACTTATGCTGGGGTAAAAATCAACGAGATGTGTATGAAA	3653
QY	3402	CAGGCGGTGGCCCTTGAGTCAATCCGAAACCTTGACAGAGAGTGAACAATGGAAGCAT	3461
Db	3654	CAGGCTATCTCCCTGTGCTCGTGGAAAACCCCTTACAGACAGTGAACCTGATGGAAGCACCC	3713
QY	3462	CCCTTAAGCAGATTCAGACACTGCAGATGGGACACACAGATGAAACAATTGACGCGCAG	3521
Db	3714	CCCGTATGCCGACTTTGACGACACAGGACCAACCCAGAAAAAGACGATTTGTGAAATTCAT	3773
QY	3522	GACATTAAGCCACTGCAGCTGTCCAGGACTCAAGTCAACAGAAAGAGGGCGGCTACT	3581

Db 3774 GAGGAGATGAGTTCGATCTGGTACCGAGTCAGGGGGCA CAGAAGCAGAGAGTTCCT 3833
Qy 3582 GCTCAGAAAGAGAGCCTTCGACACTACCTAATATGTTCCAGCCAGGAAGAACATGGG 3641
Db 3834 GCACAGAAAGAGAGCCCTCCAGCACCTTCCAGTTTGTGTTCCAGGAAGAACTAAAGAA 3893
Qy 3642 GAAGAACGAGGAGAGATGTTCTTGAACCTACACAGCAAGAGCTTACTGTGCGAGCGTG 3701
Db 3894 CAATCAAGATGGAAGACACTCTAGAGCATATACATAAAGAGGTGTCAAGTGAAGTGA 3953
Qy 3702 CCCCTTCGCAAGAGACTGAGTGGGTCAAGAGGGTGA-----GGTTGACTGGTTG 3752
Db 3954 TCCATTCTGTCAAGACTGAGGGAGCTCAAGAGGCTGACAGTATGCTGATGAGAAACC 4013
Qy 3753 GATGAGAGAAAAGTCAAGAGAAACAGGAGGTGTTGTATACCTCTGG-----ACCCAAC 3806
Db 4014 AAGACGTACCATTTTTCGAAGGACTTGAAGGGTCTATAGACACAGGCATAAACAGTCAGT 4073
Qy 3807 AGTCAAAAGGCTGTGATGTGACATATGACAGTGAAGTATGAGGAGTGGCGGTGTGAG 3866
Db 4074 CGGAAAAGGTCACTGAAGTTGCCCTTAAAGGTGAAGGGACAGAAAGCTGAATGTAAA 4133
Qy 3867 GAAAAGGAGAGTACTGAAGTGCAGAG-----TCTTAGCCTGGAGGAGGA 3911
Db 4134 AAGGATGATGCTCTTGAATCTCAGAGTACCGCTAAGTCTCTCCATCCCCGTGGAGAGA 4193
Qy 3912 GAGATGGAATCTGAGTTGAAAAGGAGAAAAGGAGAGCAAAAGCCAGAGCAAGTGAAGTGA 3971
Db 4194 GAGATGAGTGTCAAGTCGAAGAGGAGAAAACAGAGCAGAGCCAAACCCATGTGAATGAA 4253
Qy 3972 GAAGG---TGAGCAGGAAAACAGCGCTCTGTAGCATGAAGGAACTCAGGGAAGCCAGTC 4028
Db 4254 GAGAAGCTTGAGCAGCAAGACAGCTGTACCGTATCTGAAGAGGTCAAGTAAAGCAGCTCTC 4313
Qy 4029 CTGACATCTGACATGCCAGCTCAGAGAGGGGAGGACACTGGGAAGCCTTGGAGGAAGC 4088
Db 4314 CAGACGTGAATGTGCCCATCATAGATGGGCAAGGAAGTCAAGTGTGGAAAGGAAGC 4373
Qy 4089 CCTTCTC---TCCAGACCAAGACAGAAAGAGGTTGCAATAGAGGTTCAAGTTCAAGGCTG 4145
Db 4374 CCTCCTCCCTAGGTCAAGAGGAGGAGTATGCACCAAAATCAAGTTCAGAGCTCT 4433
Qy 4146 GACACACAGTCACTCAAAACAGCAGAGCTGTGAAAA-----GGTCATAGAAACCGTT 4199
Db 4434 GAGGCATCAATCTCACTAAACAGCGGCTGCAGAGGAGAAAGGCTTAGGAGAACTGCC 4493
Qy 4200 GTGATTTTCAGAGCAGGTGAAGTCCAGAGTGTGTAGGTGCACACTTATTACAGCTGAG 4259
Db 4494 AACATTTTAGAAAACAGGTGAACCGTTGGAGCCTGCAGGTGCACATTTAGTTCTGGAAGAG 4553
Qy 4260 AAGTCTCTGCAACGGGTGGCCACTGGACTCTTCAGCATGCAGAGGACACGGTACCCTG 4319
Db 4554 AATCCTCTGAATAAATGAAGACTTTTGCCGCTCATCCGGGGAAGATGCTGTGCCACA 4613
Qy 4320 GGGCCTGAGTCTCAGGAGAAATCCATCCCAATCATAGTAATCTCTGCTCCTGAAAGACC 4379
Db 4614 GGGCCGAGTCTCAGGCAAAATCGACACCACTGATAGTATCTGTACTACTACCAAGAAAGGC 4673
Qy 4380 CTACATCTGACCTACAGGAGAAATAGCCATCCAGAGAGCCGATCAGAGGAAGAG 4439
Db 4674 TTAAGTTCGACCTGGAAGGAGAGAAAACCACTCATCTGAAGTGAAGTCAAGTGAAGTC 4733
Qy 4440 GACAAGCCAGATGCTGTCTGTGCTGAGCTGACGGCAAGGAGAGTACGCAATCGAAAAAGTC 4499
Db 4734 GATCAGCAGGTGTCTGCCC-----AGGAGTCAAGTGAAGTGTAGCAATG---AGAT 4784
Qy 4500 CTCAGGCTGAACCTGAGATCCTGGAATCTGAGAGTGAAGAGCAACAAGATTGTCTGAAC 4559
Db 4785 TTAGAGCCTGAATAATGGATTGTGAACTTGAGACCAAAAGCAGTAACTTGTCCAAAC 4844
Qy 4560 GTCAATTCAGACAGCGGTGACAGTTCGCAGTAC---AGAAACAGGCCCCCGAAACTCAT 4616
|||||

Db 4845 ATCATCCAGACAGCCGTTGACCAGTTTGTACGTACAGAGAAACAGCCACCAGAAATGTTG 4904
Qy 4617 GCTTATGATTTACAGACCCAGGTTCT-----CTGCAATGCAAGGTTGACAG 4660
Db 4905 ACCTCTGAGTTACAGACACAAGCTCACGTGATAAAGCTGACAGCCAGGACGCTGGACAG 4964
Qy 4661 CAGGAGCCCAACAGATGCTGGACAAAATGAAAGATGCCAGATGAAACACCCAGTGCC 4720
Db 4965 GAAACGGAGAAAGAGAGAGGAAACCTCAGGCCCTCTGCACAGGATGAAACACCAATTTACT 5024
Qy 4721 GCAGCCCA-----GAGAGGACTTGCACAGTCTCTCAAGTCTCTGACCGTTCTGGA 4759
Db 5025 TCAGCCAAAGAGAGAGTCAGAGTCAACCGCAGTGGGACAGCACTTCTGATATTTCCAAA 5084
Qy 4760 GGCAATGGGCTCAGCCTCGGAAATGCTTGGCCGCTTTCAGTGTGAAAGCGCGGTGCAA- 4818
Db 5085 GACATGAGTGAAGCCTCAGAAAAGACCATGACTGTTTGAAGTGAAGGTCCCACTGTAAT 5144
Qy 4819 ----- 4818
Db 5145 GATCAGCAGCTGAAGAGGTGCTCTCCCATCTGAGAAAGAGGAGGTGGAGCTGGAACA 5204
Qy 4819 -----AGTAAGCATTTGAGAAAGTCCCT 4840
Db 5205 AAGTCTGTGCCAAGATGATGATGCTGCTTGTAGCAGAAAAGATAGAGAAAGTCACTA 5264
Qy 4841 CCTCAACCCCAAGA--TCCAAAAGAGAGCATGTGCTGATGGCCCTCAGCTCCAAAAGCTTA 4898
Db 5265 GTTGAACCGAAGAGATGAANAAGGTGATGATGTTGATGACCTGAAACACCAAGACTCA 5324
Qy 4899 GCCAGCAGAGGCGCAGTGCCTCTGGAACCTTAAACCAAGAAATTCGCCAGACACACAAGGA 4958
Db 5325 GCCTCGCTGATCTGATGCTCAGGAGGCTTAAACCAAGAGTCCCCAGATACAAATGGA 5384
Qy 4959 CCAAGCTAACCCAGAGGCGGATCCCCAAAAGTTGAGGTCCAGGAAGAGAAATGTCT 5018
Db 5385 CCAAAAACAAAAGAGAGGAGGATGCCAGGAATGAGAAATTGCAGGAAGGAAAAGTGCAC 5444
Qy 5019 ACCAAGTCACTCAAGAGAAACAAAGGCCAGGCAAGAGAGACCTGCAGGAGCCAAAAGGA 5078
Db 5445 AGTGAATCAGATAAAGCATCACACCCCAAGCAGAGGAGGTTACAGAAACAAAGAGAGA 5504
Qy 5079 GACCTGGCAGAACTCCTAAGATGTTAGTTG-----CTCATTTGATC 5117
Db 5505 GAATCTGCAAGTCAGAACTTACAGAACTCTTAAACATCATGCAGTTAAACTCATTTGCT 5564
Qy 5118 ATCTGTAAGACCAAGATGTGAACAAGTCACAGAACAA-----GATGCTGCTGTGGAC 5173
Db 5565 GTTGAAGACCAAGATGTGAAGCAAGTAGTAGAAGAAAATGAATGCTGCTGTGAGA- 5623
Qy 5174 CTTGAGACCAAGATTTTCAGAGCCCATGAGATCCAGAGAGCAGGGCGCTCCAATGATTTC 5233
Db 5624 CTGAAGCCAGTATTTTCAGA-ACTTTGAGNATTTGGAGAGCAGCAATCAACTGATCTCA 5682
Qy 5234 ACCAGTGAAGACCCCGACAATTTCTGAGCTTCATCGGAGCTAGAGCAGCTAACATTT 5293
Db 5683 TTTCTAGAGAGC-CCCTGACAATCTCAGGCTTCATCAGGAGCTAGAGCCATTTAACTT 5741
Qy 5294 TCCTCGTTCAAGACTGCCTTTGATTTGCCCTTGATGCGCTCGGTGATTTCTTAACATT 5353
Db 5742 TCCTCTTTCCAGACCAACCTACAAATTTTCCCTTTGATAAC----- 5781
Qy 5354 TCCTCGTTTCAAGACTGCCTTTGATTTGCCCTTTGATGCGCTCGGTGATTTTCGGATTTA 5413
Db 5782 -----CATATAAATTTGATTTA 5799
Qy 5414 AGGTCTCGCTTCTCAACCTGGAAC-CAATTTCTGCCATACCTAGTTCCCACTTCTCAAACT 5472
Db 5800 AGGTCTCAAAATCTTAACCTGGAACCTGGAGTTGGCAATACCTAGTTCTGCTTCTGAACT 5859
Qy 5473 GGAGCATCTCCTTTATGATTTATATGATGATGTTTTATGATGCTCCTCTCTCTACCTAT 5532
Db 5860 GGAGTATCATCTTTACATATTTATATGATGTTTTTAAGTAG---TCCTCCTGTATCTAT 5916
|||||

QY 5533 TGTATATTTTCTTACGTT-----AGCACAGCTTTTGTATTA 5575
 Db 5917 TGTATATTTTCTTATGTTTAAAGAAATGTGAGAGATCTACTATCTTTTGTATCA 5976
 QY 5576 TGCATATATTAACGGGTGTGACGATGAGCGCTTTGAAAGCTTCCAGGCTCACTG 5635
 Db 5977 CACAGTATATATGAGGATGTCATGTCAGGCTTGGGAGCTTTAAGCTTACGTTA 6036
 QY 5636 TAACCTGACGAAACAGATTAACATTC-----CTGGCAAGAAAGACAAAGTCTTTT 5686
 Db 6037 TATACCCACAAAAAGAGAGCTCTAGATGTAACTTCCTGATCAAGTACATTTCT 6096
 QY 5687 TAAAGTTACTGATGCTTAAATCTGTGGGCTTCTGCTCTGAAAGGTGTTTCT 5746
 Db 6097 TAAATTCATTAATGATGAGGTCCATATTTAGTGTACTCTGAAATGTGCTACTTCT 6156
 QY 5747 ATGCACAGGAGCTCAGAAATTAACCCCATTTTGAACATCCAGATGTCCTCAAT 5806
 Db 6157 ATTACACGAGAGTGTCCAAACTAAAGACATTTGAAACATACAGAAATGTTCTT 6216
 QY 5807 ACCATGATTTTCCCTTTTCTTATCCAGTCCAGTTGGAAGAGTCTCTCTG 5866
 Db 6217 ATTGGGAATTTTGC-----TTTCTAACCCAGTGGAGTTAGAAAGATTAATTTCT 6269
 QY 5867 TGTCTAGA-----TTAAGCCTGTCTCTTAATGATGACAAATGAGT 5909
 Db 6270 GGTACCAATTAATTAATTAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 6329
 QY 5910 GTGGCTTAAGCCATGAGA-----TGTTCCATAGCGAAGAAATCTGT 5953
 Db 6330 GTGCTTAATTCGAGGCAAGAGTAGATATGTTTAAATGTTAAGAAAGAAATGTT 6389
 QY 5954 TGTACGTTTTTGTGATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 6013
 Db 6390 TGTAAAG--TTTTGATTTCTACTTATATGCTGAGCTGATTCACACATGATGAATA 6447
 QY 6014 AGTCTGTTCTTAAAGATGATTTTGTATGATGATGATGATGATGATGATGATGAT 6073
 Db 6448 AGTCAGGTTCTTAAAGATGATTTTGTATGATGATGATGATGATGATGATGATGAT 6507
 QY 6074 TGCCCTCTTCTTAAAGAAAGATGATGATGATGATGATGATGATGATGATGATGAT 6133
 Db 6508 TGCCATTCCTTAAAGAAAGATGATGATGATGATGATGATGATGATGATGATGAT 6567
 QY 6134 CTGATTTAAATAAA 6147
 Db 6568 CTGATTTAAATAAA 6581
 RESULT 11
 HSU81607 6608 bp mRNA linear PRI 26-JUN-1997
 LOCUS HSU81607
 DEFINITION Homo sapiens gravin mRNA, complete cds.
 ACCESSION U81607
 VERSION U81607.1 GI:2218076
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE
 1 (bases 1 to 6608)
 Naureit, J.B., Klauck, T.M., Langeberg, L.K. and Scott, J.D.
 Gravin, an autoantigen recognized by serum from myasthenia gravis
 patients, is a kinase scaffold protein
 Curr. Biol. 7 (1), 52-62 (1997)
 JOURNAL
 MEDLINE
 PUBMED
 97153077
 9000000
 2 (bases 1 to 6608)
 Naureit, J.B., Klauck, T.K., Langeberg, L.K. and Scott, J.D.
 Direct Submission
 Submitted (10-DEC-1996) Biochemistry, OHSU, 3181 SW Sam Jackson
 Parkway, Portland, OR 97201, USA

COMMENT On Jun 26, 1997 this sequence version replaced gi:1850362.
 FEATURES
 source
 1..6608
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 192..5537
 /note="kinase scaffold protein"
 /codon_start=1
 /product="gravin"
 /protein_id="AAC51366.1"
 /db_xref="GI:2218077"
 /translation="MGAGSTTEQRSPPEQPEGSSTPAEPEPSGGSPSAAAPPTADP
 AIAAPSLATLQKNGSLSTINGVAODELSLOEDLNQKALNGALNSQEEV
 IYVNGORDSESVKSRDSEKEMATKSAVVDITDQGEENRNLEI PSESNELELLO
 PTESQADIGFKVYKRPVGRKPTVKKDKTEPTPTVLLYKDEGGGAGAGHOPS
 LGGEAASKESEPKQSTERPEETLKQESHAETSPAESQAAVEKEDEKEDEKDEK
 SKASBPSPVSTSETGTRFKPPTQGMAMRKRTSRKPEDEVEASEKKKEDEPKV
 DTEDEKAEVASSEKLTASDAQHPOEPAEASHEERLSAEKEVELPEEIVSGSGQSE
 EKPAPLATEVDEKEIVHOEVVAEYVETVEERTEOKTEVETAGSVPAEELVGM
 AEPOEAPPAKEIVKLETKETCVSGEDPTGADLSPDEKVLSPKPGVSVEMLSQRM
 KVQGSPLKLTSTGLKLSKKOKKRGGDESEBHVOVPADSDSOEOKGESA
 SPSEPEEITCLEGLAEVQDGEAEAGTSDSEKREGVTPASFRTKMTVPKKVR
 PSESDKEDELDKYSATLSTESTASBOMEAMGSVEPEPKRPVTSWEAL
 CVGSSKKRARSSDEEGPKAMGDDHQADAGDKETGDIILAGSEHDPCGS
 SPSEQASPTGEGSVTWSFKLVTPRKKSKELEKSEDSIAGSGVSTDTDEG
 KESWVSIKKFTIGRRKKRPDQKQDAPVEDAPGANGEDSDVPVAVPLSEYDAVER
 EKHEAQOQKAEOPPOKATVYSKESBSQVHMMAAAYADGTAATVILERSPSMIS
 ASVTEPLEQVEAALALTEVELEVEYIAEEEPVTEPEPENEAAGDVTSAEELP
 EAVTAETAPLSEBETGASAEETEMWASVSLDSDPTEEATPVQVEGVGD
 IAEQERTQEVLAVAEKVESQLPQTGGPEVLPVQVPAEAREPEQOAEAGLKE
 TDVVLKVDQAEATPEPTQGVGQTPSPSEFEKAPVOTESIESELVTTQCAATLKV
 KSEMMEQAIIPDSVETPTDSDSTPADADGATQKDEIVIHENEVAASGO
 SGTETAEVNAQKERPPAPSPVPOEETKQSKMEPTLETTDEKVEYETISLKTDE
 TQADQVADKTXDVPPEBGLBSIDITIVSEKYTEVALKEGTEBEAKKODALE
 LQSHAKSPSPVPEREMVQVEREKTAEPHVAEKELEHETAVTEVEYSKLQTVN
 VPLIDAKEVSLSESPPCLOEAEVCTQVQSSASFTLTAAREEVLSEETNG
 LEITGETLEPGALHLEEKSEKNEPAHIGSDAVPTGPDCAKSPVPAATTKG
 LSSDLBEKTTSLKMSDVEDVOYAOQVSVSAIDLEPENLILEETSLQNVNI
 IQTAVDQFVTEETATBMLSELQTOAHVTKASDSQAGTTEKEBEPPASADQETI
 TSKARESESTAVQASHDISKMSSESEKMTVEVSGSTVNDQLEEVLPSEBEGG
 AGTGSVEDEDDHALAERIEKSLVEPEDEKGDVDDPENONSAALADTDSGGITXSS
 PDTNGPFOKKEDEAOEVELOEGKVHESDVAITPQAOEILQKERESEAKSELTES"
 BASE COUNT 2091 a 1418 c 1827 g 1272 t
 ORIGIN
 Query Match 34.8%; Score 2144.4; DB 9; Length 6608;
 Best Local Similarity 64.3%; Pred. No. 0;
 Matches 4215; Conservative 0; Mismatches 1796; Indels 543; Gaps 45;

QY 309 GAGTCAGAGATGTGAGAGAAAGACCGAGTTGAAGAAATGGCGCCAACTCCACAGCT 368
Db 522 GACTCTGAAGATGTGAGGAAAGAGACTCCGATAAAGAGATGGCTACTAAAGTCAGCGGTT 581
QY 369 GTTGAAGATATCAAAAGGATGGCAGGAGAGACATCAGAAATAATTAATGAACAGATCCCT 428
Db 582 GTTCACGACATCACAGATGATGGCAGGAGGAGA--ACCGAATATCGAACAGATTCCT 638
QY 429 GCTTCAGAAACAATGTGAAGAAATGGTACAGCTGTGAGTCCAGGGCTAATGATGTT 488
Db 639 TCTTCAGAAAGCAATTTAGAAAGAGCTAACACAACCCACTGAGTCCAGGGCTAATGATATT 698
QY 489 GGCCTCAAGAAAGATTTAAATTTGTTGTTTAAATTCACGGTGAAGAGGATAAATAAT 548
Db 699 GGATTTAAGAAAGTGTTTAAGTTTGTGGCTTTAAATTTCACTGTGAAGAAAGGATAAGACA 758
QY 549 GAAAGTCAGATACTCTCAACTACTCACTGTCAAGAGGATGAAGCGAAGGGGAGAA 608
Db 759 GAGAAAGCTGACACTGTCCAGCTACTCACTGTGAAGAAAGATGAAGGGGAGGAGCAGCA 818
QY 609 GCCTCTGTGGAGCTGGAGACCAACAGAGCCAGTGTGAGACTGCGTCTGGAGATCA 668
Db 819 -----GGGGCTGGCGACACACAGGACCCCAAGCCTTGGGGCT-----GGAGAAGCA 863
QY 669 GCATCCAAAGAAAGTGAAGCAATCCACAGAGCAAGAGGACCCCTGAAGCAA 728
Db 864 GCATCCAAAGAAAGCAAGCAATCTACAGAGAAACCCGAAGAGACCCCTGAAGCGT 923
QY 729 GAAACAGAGCAGCAGAAATCCCCCTTCAAGCCGAATCTGATCAAGCGGCTGAGGAAGAA 788
Db 924 GAGCAAGCCACGAGAAATTTCTCCCCCAGCCGAATCTGGCCAAAGCAGT----GGAGGAA 980
QY 789 GCCAAAGATGAAGAGAGAAAGAAACAGAGAAAGAGCCCAAGTCCCAGAAATCCCCG 848
Db 981 TGCAAGAGAGGAGAGAGAGAAACAAAGAAAGAAACAACTAGCAAGTCTGCAGAAATCTCCG 1040
QY 849 AGCAGCCAGTCAACAGTGAGAGCAACATCTTCTTCAAGAGTCTTCACTACAGGTGG 908
Db 1041 ACTAGTCCCGTGACAGTGAAGCAAGGATCAACCTTCAAAAATTTCTTCACTCAAGGTGG 1100
QY 909 GCCGGCTGGCGCAAGAGACAGCTTCAAGAAATCAAAAGAGGATGATCTGGAACTGCC 968
Db 1101 GCCGGCTGGCGCAAAAGACAGTTTCAAGAACGCGAAGGAGATGAAGTGAAGCTTCA 1160
QY 969 GAGAAGAAAGGAGCAGAGGCGAGAAAGAAAGTAGCAGGAGAAAGAAAGGAAAGACAGAG 1028
Db 1161 GAGAAGAAAGGAAACAGAGCCAGAGAAAGTAGACACAGAGAAAGACGGAAGGCGAGAG 1220
QY 1029 CCAGCTCGGAGG-----GAGAGCCCGCAGAGAC 1061
Db 1221 GTTGCCTCCGAGAACTTGACCGCTCCGAGCAAGCCACCCACAGGAGCCGCGAGAAAGT 1280
QY 1062 ACAGACCAGGCGAGTTGTGACAGAGACTACGAGAAAGTGAGGCTGCCTTTGGAGAGCCAG 1121
Db 1281 GCCCAGAGCCCGGTTATCAGCTGAATATGAAGAAAGTTGAGCTGCCCTCAGAGAGCA 1340
QY 1122 GTTGGTCACTGGAGGATCGTCAGAGGAGAAAGTGTCTCTCTTTGGCAACGGAAGTGT 1181
Db 1341 GTCAGTGGCTCGCAGGACCTTCTGAAGAGAAACCTTCTCGTGGGACAGAAAGTGT 1400
QY 1182 GATGAGAAAGTGGAGCCCAACAGAA---GTTGTCAGAGGTCACGTCAGACCCGTCG 1238
Db 1401 GATGAGAAATAGAAAGTCCACCAAGAGAGGTTGTGGCCGAAAGTCCACGTCAGCACCGTG 1460
QY 1239 GAGAAGACAGAGGAGGAGGAGGAGGAGGAGGCTGAGGGGCGTGTGTTAGAA 1298
Db 1461 GAGGAGAAACCGAGAGCAGAA-----ACGGAGGTGGAA 1496
QY 1299 GGAAACAGGAGAACTCTTTGCCCTTGAGAACTGGCTGAGCCCGCAGGAGTCCCCAGGAA 1358
Db 1497 GAAACAGCAGGCTGTGTGCCAGCTGAAGATTTGGTTGGAATGGATGACAGAACCTCAGAA 1556
QY 1359 GCTGAGCCTGCTGAGGAGCTGATGAAGAGCAGAGAGATGTGTCTCTGAGGAGACCCAC 1418

Db 1557 GCCGAACCTGCCAAGGAGCTGGTGAAGCTCAAGAAACGTTGTGTTTCCGGAGAGGCCCT 1616
QY 1419 ACTCAACTGACAGACACCTTAAGTCTCTGAAGAGAACGCTGCCAAACACCCAGAGAGCAAT 1478
Db 1617 ACACAGGAGCTGACCTCAGTCTGATGAGAAGTGTGTCTCAAAACCCCGGAGGCGTT 1676
QY 1479 GTCAGTGAGTGTGAGATGTCTCTCTCAGGAAAGAAATCAAGTACAGGAAAGTCCCTTG 1538
Db 1677 GTGAGTGAAGTGAATAATGTCTCATCAGGAGAGAAATGAAGTGCAGGAGTCCACTA 1736
QY 1539 ACGAAACTCTTCACTAGTCTCAGGCTTAAAGAGCTCTCTGGGAGAGCAGAGAGGAGAA 1598
Db 1737 AAGAAGCTTTTACCAGCACTGGCTTAAAGAGCTTCTTGAAGAAAGAAACAGAGAGGAGAA 1796
QY 1599 CGAGGAGTGGGGGAGACGAAAGAGCTTGAGAAATACCAACACATTTCAACCGAATCCCCA 1658
Db 1797 AGAGGA--GGAGAGACGAGGAATCAGGGGAGCACACTCAGGTTCCAGCCGATTTCTCG 1853
QY 1659 GAGAGTGTGATGAGCAGAAAGGAGAGAGCTCTGCTGTCTCTCCCGAGGAGCTGAGGAG 1718
Db 1854 GACAGCCAGGAGGAGCAAAAGGCGAGAGCTCTGCCTCATCTCCCTGAGGAGCCGAGGAG 1913
QY 1719 ACCACGTGTCTGAGAAAGGGCGCTGGAAGCACCCAGAGTGGGAGAGCTGAGGAGGA 1778
Db 1914 ATCACGTGTCTGAAAGGGCTTTAGCCGAGGTGCACAGAGTGGGAGAGCTGAAGAAGGA 1973
QY 1779 ACTACTTCCGATCGAGAGAAAGAGAGAGGAGTCACTCTCCCTGGGCTCTCTTCAAAAAG 1838
Db 1974 GCTACTTCCGATCGAGAGAAAGAGAGAGGAGTGTCACTCTCTGGGCTCATTTCAAAAAG 2033
QY 1839 ATCGTCAACCCCAAGAAACGGGTCCGAAAGCTTTCTGAGAGTCAAGGAGGAGAGCTG 1898
Db 2034 ATGGTGACGCCCAAGAAAGCGTGTAGACGSCCTTCGAAAGTGTAAAGAAAGATGAGCTG 2093
QY 1899 GAGAAAGTCAAGAGCCCACTTGTCTCTCACTGATAGCACAGTGTCAAGAAATGCAAGAT 1958
Db 2094 GACAAAGTCAAGAGCGCTACCTTGTCTTCCACGAGAGCACAGCCTCTGAAATGCAAGAA 2153
QY 1959 GAAGTCAAAAATCTTGGTGAAGAAACAAAGCCAGAGAGAAACAAAGCGTAGGGTGGATCT 2018
Db 2154 GAATGAAGGGAGCTGGAAGAGCCAAAGCCGAGAGAACAAAGCCGAGGTGGATACC 2213
QY 2019 TCAGTGTCTTGGAGACCTGATTTTGTGTGGATCATCCAAGAGAGAGACAGAGAGGCA 2078
Db 2214 TCAGTATCTGGGAAGCTTTAATTTGTGTGGATCATCCAAGAAAGAGACAGAGAGAGG 2273
QY 2079 TCCTCTTCAGATGATGAAGAGGGCCCAAGGACACTGGGAGGGGACAGTCAAGAGCAGAG 2138
Db 2274 TCCTCTTCTGATGAGGAAGGGGACCAAAAGCAATGGGAGGAGACCACAGAAAGCTGAT 2333
QY 2139 GAGGCCAGCAAGACAAAGAGCCGGAACAGAGCTGTTCTCCAGCAGCACCAGGAGCAG 2198
Db 2334 GAGGCCGGAAGAAAGACAAAGAGACGGGGGACAGCGGATCTTGTCTGTTCCCAAGAACAT 2393
QY 2199 GACCAAGCCGCAAGAAAGTTCCTCACCCGAGCAGCGGGAGAGCCCTTCCGAAGGGGAGGT 2258
Db 2394 GATCCAGGCGAGGAGTTCCTCCCGGAGCAGCTGGAAGCCCTACCGAGGGGAGGC 2453
QY 2259 GTCTCCACTTTGGAGTCATTTAAAGATTTAGTCACTTCAAGAAAGAAATTCAGAGTCAAAA 2318
Db 2454 GTTTCACCTCTGGAGTCATTTAAAGGTTAGTCAAGCAGGAGGAGGAGGAGTCCCAAG 2513
QY 2319 CTGGAAGAGAAAGCCGAAGAC-----TCTAGTGTAGAGCAGTGTCTCACTGAG 2366
Db 2514 CTGGAAGAGAAAGCCGAAGACTCCATAGCTGGGTCTGGGTGAGAACATTCCTCAGAG 2573
QY 2367 ATCGAAACCGAGTAGAGAGAAATCTTGGGTTTCCATTTAAGAAATTCATCCCGGAGCGCG 2426
Db 2574 ACTGAACCCGGTAAAGAGAAATCCTGGGTCTCAATCAAGAGTTTATTCTTGAGCAGAGG 2633
QY 2427 AGAAAGGGCGAGACGGGAGGAGAGCAAGGCCACTGTGTGAAGACTCAGGGCCAGTGGAG 2486

Db 2634 AAGAAAAGCCAGATGGGAACAAGAACAGGCCCTGTTGAAGACGAGGGCCAAACAGGG 2693
Qy 2487 ATAAATGAGACGACCTTAATGTCCAGCCGTCTGTCTGTGATATAATGCACTG 2546
Db 2694 GCCAACGAAATGACTGTGATGTCCCGCCGTGTGCTCTGTCTGATGATGATGCTGTA 2753
Qy 2547 GAGAGGAGGAAGATGAG-----AGCCAGGGGAAATAGCGAGCTCCAGCTGTG 2597
Db 2754 GAAAGGGGAAATGAGGACACAGCAAGCCAAAGGCGCAGAGCAGCCGAGCAGAAAG 2813
Qy 2598 GGGCTGTGATCTGTCCGAGAGGCTCAGTAAGACTGTGTCCACACTGTGATGTGCA 2657
Db 2814 GCAGCACTGAGGTGTCCAGGAGCTCAGCAGAGTCAAGTTATGATGAGCAGCAGCT 2873
Qy 2658 GTCAATTGAGGACCAAGGCGAGTCAACAGTGTCAAGAGCCGTTCTCTTCTGTGATATCC 2717
Db 2874 GTGCTGACGGGACGAGGGGAGCTACCATTAATGAAAGAAAGGTCTCTTCTGTGATATCT 2933
Qy 2718 GCTTCCCTTAACAGAACCTTGAACACACAGCGGAGAAAGCCATGSCACTGTGAGAG 2777
Db 2934 GCTTATGACAGAACCTTGAACAGTAAGAGCTGAAGCCGCACTGTTAACTGAGAG 2993
Qy 2778 GTCACTGAAAAAGACATCATGTC---AGAAAGAACTCTGTGTCTCACCCAGAGCTTTACCA 2834
Db 2994 GTATTGAAAAGAGAGTAATTGCAAGAAAGAAACCCCAAGGTTACTGAACCTCTGCCA 3053
Qy 2835 GAGGTTAAAGATGCCCATGACGACATGTGTCAACAGTGAAGTGAATTTCACTCAGAGCT 2894
Db 3054 GAGAACAGAGAGGCCCGGGGCGACACGGTCTTAAGTGAAGCGAAATGACCCCAAGCT 3113
Qy 2895 GTGACAGCACAAGACCTCAGAGGCTCCGTACTGAAGAAAGTTACCGAAGCATCGGG 2954
Db 3114 GTGACAGCTGCGAAGAACTGCAAGGCCATTTGGTTCGAAAGAAAGAACGAAACACTGTGT 3173
Qy 2955 GCCAAGAGACCAACAGACATGTGTCCGCAAGTTTCCAGCTGACTGATCTCCCAAGACC 3014
Db 3174 GCTGAAGAGACCAACAGAAATGAGTGTACAGATCTCCAGTTAAACGACTCCCAAGACACC 3233
Qy 3015 ACAGAGAAACCAACCCCACTTCAAGAGTGAAGAGTGTGTCTAGATACAGAAAGAG 3074
Db 3234 ACAGAGAGGCCACTCCGCTGCAAGGAGTGAAGGAGTGAGTGTACCTGAATGAAGAGCAA 3293
Qy 3075 GAGCCCAAGACCAAGGCCATCTCTCAAGCCGTTGACAGCAAGAGTGAAGAGAGTCCAG 3134
Db 3294 GAGAGCGGACCTCAAGAGGTCTCTCAAGGCAGTGCAAGAAAGTGAAGAAATCCAG 3353
Qy 3135 GTGCCCTGCAACC---CAGACTGTGACAGAAACGGGGTCAAAAGCACTGAGAAAGTTGAG 3191
Db 3354 CTGCTGTGCAACCGGTGGGCAAGAAATGTGCTTCAAGCTGTGACAGAGCAGAGGAGAA 3413
Qy 3192 GAGGTAGAGAGGACTCCGAAGTGTGCTTGGAGAAAGAGAGAGCTTATGCCGAAA 3251
Db 3414 AACACGAAAGACAGGCTGAAGGTGGGTCTGAAGAAAGAGACGATGTAGTGTGAAA 3473
Qy 3252 GGAACCGTGCAGAAAGTGAAGTGCATCTTGAACAGGGCTCTGAGCTGAGAGGCT 3311
Db 3474 GTAGATGCTCAGAGGCAAAAACCTGAGCTTTTACACAGAGGAGGTGTGGGGCAGACC 3533
Qy 3312 ACTCCAGAGAGCCTT---AAGTTCTGAAGTCAACGGCAGATGTAGAC-----CAT 3359
Db 3534 ACCCCGAAAGCTTTGAAAAAGTCTCTCAAGTCAACAGAGCAGTAAGTCCAGTAGCTT 3593
Qy 3360 GTGCGCAAGTCCA-----GTTATCAAGTCCAAGCAGCTGATGAA 3401
Db 3594 GTAAACCACTTGTCAAGCCGAAACCTTAGCTGGGGTAAAAATCAGAGAGATGTGATGAA 3653
Qy 3402 CAGGCGGTGGCCCTGAGTCACTCCGAAACCTTGAACAGACAGTGAACAATGGAAGCACT 3461
Db 3654 CAGGCTATCTCCCTCTGACTCTGAGTGAACCCCTTACAGACAGTGAAGTGAAGACACC 3713
Qy 3462 CCCTTAGAGATTCAGACACTGAGATGGGACACAGCAAGATGAACCAATTGACAGCCAG 3521
Db 3714 CCGTAGCCGACTTTGACGACAGGACCAACCCAGAAAGACGAGATGTGTGAAATCCAT 3773

Qy 3522 GACAGTAAAGCCCACTGAGCTGTCAAGGCACTCAAGGTCAAGAAAGAGGCGGCTACT 3581
Db 3774 GAGAGAAATAGAGTCCGATCTGGTATCCCACTCAGGGGGCACAAGAGAGGCAAGTTCT 3833
Qy 3582 GCTCAGAAAGAGAGCCTTGCACACTAAGTAATGTTCCAGCCCAAGAAACATGG 3641
Db 3834 GCACAGAAAGAGAGCCTCCAGACCTTCCAGTTTGTGTTCAGAGAAAGAACTAAAGAA 3893
Qy 3642 GAAGAACACAGAGAGATGTTCTTGAACCTTACACAGAAAGCTTCTGTGACCGCTG 3701
Db 3894 CAATCAAAAGATGAGAGCACTTGAAGCATACAGATTAAGAGGTGATGAGAACTGTA 3953
Qy 3702 CCGTTCTGCAAAAGACTGAGTGGGTCAAGAGGTGA-----GTTTACTGGTGG 3752
Db 3954 TCCATTCTGTCAAAAGACTGAGGGGACTCAAGAGGCTGACAGATGTGATGAGAAACC 4013
Qy 3753 GATGAGAAAAAGTCAAAAGAAACAGAGGTGTTGTAACCTGTG-----ACCAAC 3806
Db 4014 AAAGACGTACATTTTTCGAAGGACTTGAGGGGTCTATAGACACAGGCATTAACAGTCAGT 4073
Qy 3807 AGTCAAAAGGCTGTGATGTGACATATGACATGTAAGTGAAGTGGAGCGGGTGTGAG 3866
Db 4074 CGGAAAAAGTCACTGAAGTTGCCCTTAAAGGTGAAGGAGCAGAAAGAGTGAATGTAAA 4133
Qy 3867 GAAAGAGAGATCTGAAGTGTGACAG-----TCTTACCTGAGAGAGGA 3911
Db 4134 AAGATGATGTCTTGAACCTGACAGATCACGCTAAGTCTCTCCATCCCGGTGAGAGA 4193
Qy 3912 GAGATGAAACTGACGTTGAAAGAGAAAGGAGACAAAGCCAGACAGTGAAGTGA 3971
Db 4194 GAGATGTGATTCAGTGTGAAGGAGAAACAGAAACAGAGGCCAACCATGTGTAAGAA 4253
Qy 3972 GAAGS---TGAGCAGGAAACAGCGGCTCTGAGACATGAAAGAACCTTACGGGAACCACTC 4028
Db 4254 GAGAACTTGAACAGAAACAGCTGTTAACGTATCTGAAGGTCAAGTAAACACTCTCTC 4313
Qy 4029 CTGACACTTGAACATGCCAGTCTCAGAGAGGGGAGGACACTGAGAACCTTGAAGAGC 4088
Db 4314 CAGACAGTGAATGTGCCATATAGATGGGCAAGAGAGTCAACAGATTTGAAAGAGAC 4373
Qy 4089 CCTTCTC---TCCAGACCAAGAACAGAGGTTGATAGAGTTCAAGTTCAAGCTG 4145
Db 4374 CTTCTCTCTGAGTCAAGTCAAGAGGAGGATATGACCAAAATTCAGAGTCTCT 4433
Qy 4146 GACACACATGATCTCAAAACAGCAGAGCTGTGAAAA-----GTCATGAAACGTT 4199
Db 4434 GAGGATCATTCATCTTAAACAGCGGCTGACAGAGGAGGAAAAAGTCTTAAAGGAAACTGCCC 4493
Qy 4200 GTGATTTCAAGAGCAGGTGAAGTTCAGAGTGTGATGAGTGAACCTTATTCAGCTGAG 4259
Db 4494 AACATTTTGAAGACAGTGAACGTTGAGGCTGACAGGTGTGACATTTAGTCTGGAAGAG 4553
Qy 4260 AAGTCTCTGCAACGGGTGGCACTGACTCTTCAAGCATGACAGAGACAGCGTACCCCTG 4319
Db 4554 AATCTCTGAAAAAATGAAAGACTTGGCCCTCATCCAGGGGAAAGATGTGTGCCACA 4613
Qy 4320 GGGCTGTAGTCTCAGGCGAATTCATCCCATCATAGTAACTCTCTGCTCTGAAAGCAC 4379
Db 4614 GGGCCGAGCTGTCAAGCAAAATTCGACACAGTGAATGATGTGCTACTTCAACAAAGAGGC 4673
Qy 4380 CTACATCTGACCTTACAGAGAAATTAAGCGCATCCACAGAGAGAGCATCAGAGAGAG 4439
Db 4674 TTAAGTTCGACCTGAAAGAGAGAAACACATCATCTGAAGTGAAGATCAGATGAATGC 4733
Qy 4440 GACAAAGCAGATGTGTCTGATGTGACGAGCAAGAGATGACGAAATGAAAAAGTCA 4499
Db 4734 GATGAGCAGGTGTCTTGGC-----AGAGGTCAAAAGTGAAGTGTGCAATGTG---AGGAT 4784
Qy 4500 CTCAAGGCTGAACCTGAGATCTGTGAACCTTGAAGATGAAGACAAATGTGTGTGAC 4559
Db 4785 TTAGAGCTGAAAAATGGGATTTTGAAGCTTGAAGCAAAAGACAGTAACTTGTCCAAAAAC 4844

QY 4560 GTCAATCAGACGCGTTGACCGTTCGACGTTAC---AGAAACAGCCCGGAACTCAT 4616
DB 4845 ATCAATCAGACGCGTTGACCGTTCGACGTTACGACGAAAGAAACCGCCGAAATGTTG 4904
QY 4617 GCTTATGATTCACAGACCCAGGTTTC-----CTGCATGCGAGGCTTGACAG 4660
DB 4905 ACCTGAGTTACAGACACAACTCACTGATAAAGCTGACAGCCAGGAGCTGACAG 4964
QY 4661 CAGGAGCCCAACGATGCTGGACAAAAATGAAAGATGCAAGATGAAACACCCAGTGCC 4720
DB 4965 GAAACGAGAGAAAGAGAGGAAACCTCAGGCTCTGCAAGAGTGAACACCAATTAAT 5024
QY 4721 GCAGCCCA-----GAGAGGACTTCAAGTCTGACCGTTCGGA 4759
DB 5025 TCAGCCAAAGAGAGTCAAGTCAACCGCAGTGGGCAACATCTGATATTTCCAAA 5084
QY 4760 GGATGGGCTCAGCCTCGGAAATGCTTGGCGGCTTCAGTGTAAAGCGCGGTGTCAA- 4818
DB 5085 GACATGAGTGAAGCTCAGAAAGACCATGCTGTTGAGGTAGAGGTTCCACTGTAAAT 5144
QY 4819 ----- 4818
DB 5145 GATCAGCAGCTGGAAGAGTGTCTCTCCATCTGAGGAAGGAGGCTGGAGCTGGAACA 5204
QY 4819 -----AGTAAGCATTTGAGAGCTGCCT 4840
DB 5205 AAGTCTGTGCGCAGAGATGATGTCATGCTTGTAGCAAGAAAGATGAGAGATCACTA 5264
QY 4841 CCTCAACCCCAAGA--TCCAAAAGGAGCATGCTGCTGATGGCCCTCAGCTCCAAAGCTTA 4898
DB 5265 GTTGAACCGAAGAGATGAAAGGATGATGTTGATGACCTGAAACCCAGAACTCA 5324
QY 4899 GCCAGGACAGGCGCATGCTCTGGAACCTTAACCAAGAAATCCCCAGACACCAAGGGA 4958
DB 5325 GCCCTGGCTGATCTGATGCTCAGGAGGCTTAACCAAGAGCTCCCCAGATACAAATGGA 5384
QY 4959 CCAAGCTAACGAGAGGCGGATCCCCAAAGTTGAGTCCAGAGAAAGAAATGTCT 5018
DB 5385 CCAAAACAAAAGAGAGAGGAGTGGCCAGGAAATGAGATTCAGGAGGAAAGTGCAC 5444
QY 5019 ACCAAGTCAGTCAAGAGAACAGGCGCCAGGAGAGGAGGAGCTGAGGAGCCAAAGGGA 5078
DB 5445 AGTGAATCAGATAAGCGATCACACCCCAAGCAGAGGAGTTACGAAACAGAGAGA 5504
QY 5079 GACTGCGCAGAACTCTAAGATGTTAGTTG-----CTCAITTTGATC 5117
DB 5505 GAATCTGCAAGTCAGAACTTACAGAACTTTAAACATCATGCTGAGTTAACTCATTTGCT 5564
QY 5118 ATCTGTAAGACCAAGATGTGAAACAAAGTCACAGAACAA-----GATGCTGCTGTTGGAC 5173
DB 5565 GTTTGGAAGACCAAGATGTGAAGACAAAGTAGTAGAGAAATGAATGCTGCTGCTGAGA- 5623
QY 5174 CTTGAGACCAAGATTTTCAGAGCCATGAGATCCAGAGAGCAGGCGCTCCAAATGATTC 5233
DB 5624 CTGAAGACCAAGATTTTCAG-ACCTTGAGAAATTTGAGAGCAGGACATCAACTGATCTCA 5682
QY 5234 ACCCAGTAGAGACCCCGCAATTTCTGAGGCTTCATCGGAGCTAGAGCCAGCTAACAT 5293
DB 5683 TTTCTAGAGAGC-CCCTGCAATCTGAGGCTTCATCAGAGCTAGAGCCATTAACAT 5741
QY 5294 TCCTCGTTTCAAGACTGCTTTGATTTGCCCCCTTGATGCGCTCGGTATTTCTTAACAT 5353
DB 5742 TCCTCTTTTCAAGACCAACCTACAAATTTTCCCTTGATAAC----- 5781
QY 5354 TCCTCGTTTCAAGACTGCTTTGATTTGCCCCCTTGATGCGCTCGGTATTTCTCGAATTTA 5413
DB 5782 -----CATATAAATTTCTGAATTTA 5799
QY 5414 AGGTCTGCTGCTTCAACCTGGAAAC-CAATTTCTGCCATACCTAGTTCCACTTCTCAAACT 5472
DB 5800 AGGTCTTAATTTCTTAACCTGGAACTGGAGTTGGCAATACCTAGTTCTGCTTCTGAACT 5859
QY 5473 GGAGCATCTCTTATGATTTATATGATGTTTTATGATGTTTATGATGTCCTCTCTGACCTAT 5532

DB 5860 GGAGTATCATCTTTTACATATTATATGATGTTTTAAGTAG---TCCTCTCTGTATCTAT 5916
QY 5533 TGTATATTTTTTTTCTAAACGTT-----AAGCACATGCTTTTTGTATTA 5575
DB 5917 TGTATATTTTTTTTCTAAATGTTTAAAGAAATGTGACAGGATCTACATGCTTTTTGTATCA 5976
QY 5576 TGCATATATAACGGGTGTGACGCCATAGCAGCTTTTGAAGCTCCAAAGCTCAACTG 5635
DB 5977 CACAGTATATGATGGGCATGTGCCATAGTCAGGCTTGGGAGCTTTAAGCCTCAGTTA 6036
QY 5636 TAACCTGCGAGCAACAGATAACATTC-----CTGCAAGAAAGAGACAAGTCTTTTT 5686
DB 6037 TATAACCCCAAAAAACAGAGCCTCTAGATGTAAATCTCTGATCAAGGTACAATTTCT 6096
QY 5687 TAAAGTTTACTGATGCTTAGATCTGCGGCTTCTAGTCTCTGAAAGTGGTGTCTTTCT 5746
DB 6097 TAAATTTCTAATGATGAGGTCCATATTTAGTGTACTCTGAAATTTGGTCACTTTCT 6156
QY 5747 ATGCACAGCGAGCTCAGAAATAAAACCCCATTTTGAACATCCAGGATGTCCCAATAT 5806
DB 6157 ATTACACGAGGTGTGCAAAACCTAAAAAGCATTTTGAACATACAGAAATGTTCTATTGTC 6216
QY 5807 ACCATGATTTTTTCCCTCTTTTGTCTAATCCAGTCCAGTTGGAAAGAGTCTCTCTG 5866
DB 6217 ATTGGGAAATTTTGC-----TTTCTAAACCCAGTGGAGGTTAGAAAGAAATTTATTTCT 6269
QY 5867 TGTGACA-----TTAAGCCCTGCTCTTAATGATATGACAAATAGT 5909
DB 6270 GGTAGCAAAATTAATTTTACATCTTTTCTACTGTTAGTTGTTGGACCGATAAGT 6329
QY 5910 GTGCTTAAGGCCATGAGA-----TGTTTCTTAATGCAGAAAGAAATCTGT 5953
DB 6330 GTGCTTAATCTCGAGCAAGTAGTAAATATGTTTATATGTTATGAAGAAAGAAATGT 6389
QY 5954 TGTACGTTTTTTTGTATGTAATCTCTTATGCTGAGCCAAATTCATATGAGATGCAAGTG 6013
DB 6390 TGTAAAG--TTTTGATTTCTACTCTTATATGCTGGAATTCACATGCGCATGAAATA 6447
QY 6014 AGTCCTGTTCTTTACAGATGTTATTTGATAGACTGAGTTGTTCTGTTGTTATATCTG 6073
DB 6448 AGTCAGGTTCTTTACAAATGGTATTTTGTATGATCTGGAATGTTGTGCGCATATTTG 6507
QY 6074 TGCCCTCTTTTAAAGAACAAATGTCATTTATGTTCTTTGGATAAATTTGATTTGACAA 6133
DB 6508 TGCCATCTCTTTTAAAGAACAAATGTTGCAACACATTTGATTAAGTTCTGATTGACGA 6567
QY 6134 CTGATTTTAAATAAA 6147
DB 6568 CTGATTTTAAATAAA 6581
RESULT 12
AF001504
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

AF001504
Homo sapiens myasthenia gravis autoantigen gravin mRNA, partial cds.
AF001504
AF001504.1
GI:2160783
Homo sapiens.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 6297)
Bowditch, R.D. and Ginsberg, M.H.
Sequence of gravin cDNA isolated from a human umbilical vein endothelial cell library
Unpublished
2 (bases 1 to 6297)
Bowditch, R.D. and Ginsberg, M.H.
Direct Submission
Submitted (29-APR-1997) Biochemistry, University of Oklahoma Health

Sciences Center, PO Box 26901, BMSB 853, Oklahoma City, OK 73190,
USA

FEATURES

source

Location/Qualifiers

1. .6297
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="endothelial cell"
/tissue_type="umbilical vein"
1. .5226
/note="kinase scaffold protein"
/codon_start=1
/product="myasthenia gravis autoantigen gravin"
/protein_id="AAB58938.1"
/db_xref="gi:2160784"

CDS

translation="DPAIADSPATKLLQKNGSLTINGVADEPDELQSDGDLNGOK
ALNGOALNSOEVEEVTEVQGRSDSDVSDKDMATSAVNHITDQSETEB
IIEOIPSSNLELTOPTESQANDIGKFKVFKTKDKTKDPTDVLTK
KDEBGAAGADHDPSLAGEAASKESEPKSTKEPELTKQSHALISPPAESQ
AVECKEKEKQEKESKSAESPTSPSTSPKPKFKFGMAKSKTSFKPK
DEVAEKKKEQEPKDTBEDGKAABASEKLTASBOAHPOEPAESHEPRLA
VELSEOVSGSGSPSEKPAPLATEVEFDEKIEVQEVNAEIVSTVEERT
VEETASGVPABELVEMDAEPQAPAEVLKJKTQVSGEDPTGADLSPEKVL
PEGVSEVEMLSQERMKVQSPKLTFTSTGLKLSGKQKGRGGDESGHTV
PADSPDQOEKGSASSPSEPEITCLEGLAVOODGAEAGATSDGKEGVT
EAPRKMVTPTKRVRSRPSSEKDELTKVKSATLSTSESTASHEMKSVEPK
EAPRKMVTPTKRVRSRPSSEKDELTKVKSATLSTSESTASHEMKSVEPK
TDGILAGQEHDPQGSPPQASPTGEBVSTWESFKRLVTRKSKLSEKSE
SISAGVHSTPTDTEPGKESWVSIKKRIPRRKRPDQKQAPVEDAGTGNED
SDPAVPLSEYDAVERKMEAOQAQKSAEQPEQKAEVSEKLSAEPPVTEPL
GTRPATLIEERSPSWISAVTEPLEOVAEALTEVLAEREVLAEEPPVTEPL
NREARGLTVSBAELTPRAVTAETAGPLGABEETMAAEETMEVAVALDSD
TTEETVQVEVGVPIEDQERPTQELQVAEVEKESOLPTGSGEDPLQVOA
EAREBEQAEASGLKETDVLKVAQDAKTEPTQGVQVQDTPESEKAPQTESI
ESSELVITCOAETLAGVSOEMVMEQALPPSPVETPTDSTPVAADPAEGTOK
DEIVEIHENEVAAGTOSGTEAEAVPAKERPAPSPSFVEQEKESKDELT
DKYEVETVSTLSTKEGTQADQVADBEKTKVPPFEGLEGIDTGIIVSRKTEVAL
KGEVBEAECKKDDALEHOSHAKEPSPVEREMVQVREKTEAPTHNEKLEHT
AVTVEEVSQOLLQTVNPIIDQAEVSSLSGSPPCGQGBEAVOTKIQVONSERSPT
LTAABEEKVLGETANILLETETLEPAHAILLEKESSEVQACDQKSVAIIEB
COAKSTPIVATATTKKGLSDLEKETSLLKMSSEBQACDQKSVAIIEB
ENGILEETKSKLVQNIQTADQVRETEATETMLTSELQTAHVKAQSODAG
TEKIEEEOASQADDEPTITSAKESSESTAVQAHEDIKMSSEAEKTEVSEVST
NDQIEEVVLSSEEGGAGAGTSPVDDGHALIARIKSLVPEDEKGDVDPEN
ONSALADTDSAGLTKESEPDNTPKQKEKEDAEVLELDGKHSBDALITPQAEEL
QKQERESAKSELTES"

BASE COUNT 2051 a 1310 c 1708 g 1228 t

ORIGIN

Query Match 34.7%; Score 2136.8; DB 9; Length 6297;

Best Local Similarity 64.5%; Pred. No. 0; Mismatches 1747; Indels 525; Gaps 41;

Matches 4120; Conservative 0;

153 GACCCCGCGAGCGGAGACCCCGCCACCAAGCTCCACAGAGAATGGCAGTGTCTT 212
7 GCATGCTGCTCGGACCCCGCCACCAAGCTCTTACAGAGAAATGCTACGCTGTCCACC 66
213 GTCAACGCGTGAAGTGAAGAGATGTCATGTCACAGAGAGAGAAACCAAGAGGAGC 270
67 ATCAATGGCTAGCTAGCAAGATGAGCTCAGCTCCAGAGAGGTGACCTTAATGGCCAG 126
271 -----AGNAGNAGNAGTCTTGAAT 220
127 AAAGAGCCCTGAACGCTCAAGAGCCCTTAACAGCCAGAGAGAGAAAGAAATGCTGTC 186
291 GAGAGTGTGGAAGAGAGATGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATG 350
187 ACAGAGTGTGGAAGAGAGATGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATG 246
351 GCGGCAACTCCACAGCTGTTGAAGATATCAAAAGATGGGAGAGAGAGATCAAGAA 410
247 GCTACTAAGTCAAGCGGTGTTCACGACATCAAGATGATGGGAGAGAGAGAGAGAGAG 306
411 ATTAATTGAACAGATCCCTGCTTCAGAAAACAATGTGAGAAATGTGACAGCTGCTGAG 470

307 ATATCGAACAGATTCCTTCTCAGAAAGCAATTTGMAAGCTAACACAAACCCACTGAG 366
471 TCCAGGCTATATATGTTGGCTTCAGAAAGTATTTAAATTTGGTTTAATTACAG 530
367 TCCAGGCTATATATGTTGAATTTAAAGAGTGTTAATTTGGCTTTAAATTTACT 426
531 GTGAAGAGATATAAAATGAAGAGTCAATATCTGTCACATCTCACTGTCAAGAGAT 590
427 GTGAAAAAGATTAAGACAGAGAGAGCTGACATCTGACATCTGACATCTGAGAAAGAT 486
591 GAAGGCAAGGGGAGAGAGCTTGTGAGAGCTGAGACCAAGAGAGAGAGAGAGAGAGAG 650
487 GAAGGAGAGGAGAGAGCA-----GGGGCTGGCCAGCACCAAGAGAGAGAGAGAGAG 537
551 ACTCCGCTCGAGAGTACAGATCCCAAGAAAGTGAAGTGAAGCAATCCACAGAGAGAG 710
538 GCT-----GGAGAGAGAGAGATCCCAAGAAAGGAGAGAGAGAGAGAGAGAGAGAGAG 591
711 GAAGGCAAGCTGAG 770
592 GAAGAGAGCTGAG 651
771 CAAGCGCTGAG 830
652 CAA---GAGTGAAG 708
831 AAGTCCCAAGATCCCGAG 890
709 AAGCTGAGAGATCTCGAGATCTCGAGATCTCGAGATCTCGAGATCTCGAGATCTCGAG 768
891 TTCTTCACTCAGGTTGGGCGCGCTGGGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 950
769 TTCTTCACTCAGGTTGGGCGCGCTGGGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 828
951 GATGATCTGAGAACTGCCAG 1010
829 GATGAGTGAAG 888
1011 GAAAAAG 1043
889 GAAG 948
1044 CAGAGAGCGGAG 1103
949 CAGAGAGCGGAG 1008
1104 CTGCTTTGAG 1163
1009 CTGCTTTGAG 1068
1164 TTGGCAAG 1220
1069 TTGGCAAG 1128
1221 GTCCAGTGAAG 1280
1129 GTCCAGTGAAG 1170
1281 GGGGCGTGTGTGAG 1340
1171 -----ACGAGAGTGAAG 1224
1341 CAGAGAGTCCCCAG 1400
1225 GATGCAAG 1284
1401 GTCTCTGAG 1460
1285 GTTTCGGAG 1344
1461 AAACACCCAG 1520

Db 1345 AAACCCCGAAGCGTTGTGAGTGGAAATGCTGTCTATCACAGGAGATGAAG 1404
Qy 1521 GTACAGGAAGTCCCTTGAAGAAACTCTTCAAGTAGCTCAGGCTTAAAGAGACTCTCTGGG 1580
Db 1405 GTGACGGGAAGTCCACTAAAGAACTTTTACCAGCACTGGCTTTAAAAAAGCTTCTGGA 1464
Qy 1581 AAGAAAGCAGAAAGGGAACAGGAGAGTGGGGACACAGAGCCTGGAGATACCAACAC 1640
Db 1465 AAGAAACAGAAAGGAAGAGGA--GGAGGACACAGGAATCAGGGGACACACTCAG 1521
Qy 1641 ATTACACCGGAATCCCGAGAGAGTGTGATGAGCAGAAAGGAGAGAGCTCTGCGCTCGTCC 1700
Db 1522 GTTCCAGCCGATTCTCCGACAGCCAGGAGAGCAAAAGGGCAGAGCTCTGCTCTATCC 1581
Qy 1701 CCGAGAGCCTGAGGAGACCACTGTCTGAGAAAGGCGCGTGGAGAACCCCGAGAT 1760
Db 1582 CCTGAGGAGCCGAGGAGATCACGTGTCTGAAAAGGCTTAGCCGAGGTGCAGCAGAT 1641
Qy 1761 GGGGAAGCTGAGGAGAGACTTACTTCCGATGGAGAGAGAGAGAGAGGATCACTCCC 1820
Db 1642 GGGGAAGCTGAAGAGGAGCTACTTCCGATGGAGAGAGAGAGAGAGGAGGAGGAGTCACTCCC 1701
Qy 1821 TGGGCATCCTTTCAAAAGATGGTGACACCCCAAGAAACGGGTCCGAAGACCTTCTGAGAT 1880
Db 1702 TGGSCTTTCAATCAAAAGATGGTGACGCCCAAGAGAGCGTTAGACGGCTTCGGAAGT 1761
Qy 1881 GACAGGAGAGAGCTGAGAGAGGTCAAGAGCCCACTTGTCTCACTAGTAGACACA 1940
Db 1762 GATAAAGAGATGAGTGGACAGGTCAAGGTCAAGAGCGTACCTTGTCTTCCACCGAGACACA 1821
Qy 1941 GTGTCAAAATGCAAGATGAGTCAAACTGTTGGTGAGGAAACAAAGCCAGAGAACCA 2000
Db 1822 GCCTCTGAAATGCAAGAGAGAAATGAAGAGGAGCGTGGAAAGGCCAAAGCCGGAAGACA 1881
Qy 2001 AAGCGTAGGTGGATATCTCAGTCTCTTGGGAAGCACTGATTTGTGCGGATCATCAAG 2060
Db 1882 AAGCGCAGGTGGATATCTCAGTATCTTGGGAAGCTTAAATTTGTGTGGATCATCCAG 1941
Qy 2061 AAGAGCAAGGAGGATCTCTTTCAGATGATGAAGAGGGGCCAAGGACACTGGGAGG 2120
Db 1942 AAAAGACAGGAGAGGTCCTCTTCTGATGAGGAAGGGGACCAAAAGCAATGGGAGGA 2001
Qy 2121 GACAGTCAAGAGAGAGGAGCCAGCAAGACAAAGAACCCGGAACAGAGCTGTTCT 2180
Db 2002 GACCACAGAAAGCTGATGAGCGCGGAAAGACAAAGAGACGGGACAGACGGGATCCTT 2061
Qy 2181 GCCAGCACAGGAGCAGGACCAAGCCCAAGGAAGTTCTCACCCGAGCCAGCGGGAAGC 2240
Db 2062 GCTGTTCCCAAGACATGATCCAGGGCAGGAAAGTTCTCTCCCGAGCAAGCTGGHAGC 2121
Qy 2241 CTTTCCGAAGGGAGGTGTCTCACTTGGGAGTCAATTTAAAGATTAGTCACTTCCAAGA 2300
Db 2122 CTACCCGAAGGGAGGCGTTTCCACCTGGGAGTCAATTTAAAGTTAGTCACGCCAAGA 2181
Qy 2301 AAAAATCCAGTCAAACTGGAAGAGAAAGCCGAAGAC-----TCTAGTGTA 2348
Db 2182 AAAAATCAAGTCCAGCTGGAGAGAGAAAGCCGAAGACTCCATAGCTGGGTCTGGTGTA 2241
Qy 2349 GAGCAGTGTGCTCACTGAGATCGAACCCAGTAGAGAAATCTTGGGTTTCCATTTAAGAAA 2408
Db 2242 GAACATTCCTCACTCAGACATGAACCCGGTAAAGAGAAATCTGGGTCTCAATCAAGAG 2301
Qy 2409 TTCAATCCCGGACGGCGGAAGAAAGGGCAGACGGGAGCAAGAACCAAGCACTGTGGAA 2468
Db 2302 TTTATTCTGGACGAAGAAAGGAGCCAGATGGGAAACAGAACAAAGCCCTGTGTA 2361
Qy 2469 GACTCAGGCGAGTGGAGATAATAGAGACGACCTTAATGTCCAGCGCTGTGCTCTG 2528
Db 2362 GACGACGGGCAACAGGGGCCAACGAAGATGACTCTCATGTCTCCGGCGGTGGTCCCTCTG 2421
Qy 2529 TCTGAGTATATGACGTGGAGGGGAGAGATGG-----AAGCCGAGGGAATACG 2579
Db 2422 TCTGAGTATGATGCTGTAGAAAGGGAGAAATGGAGGCAAGCAAGCCCAAAAGCGCA 2481

Qy 2580 GAGCTGCCCCAGCTGCTGGGGGCTGTGTAGCTGCCAGGAGCTCAGTAAGACTCTGGTC 2639
Db 2482 GAGCAGCCCGAGCAGAAAGGCGAGCCACTGAGGTGTCCAAAGAGCTCAGCGAGACTCAGTT 2541
Qy 2640 CACACTGTGAGTGTCCAGCTCATTGATGGACCCAGGGCAGTCCAGTGTCCGAAGCGG 2699
Db 2542 CATATGATGGCAGAGCTGTGCTGAGCGGACGAGGGCAGTACCATTTATTGAAGAAGG 2601
Qy 2700 TCTCCTTCTGGATATCCGCTTCCGTAACAGAACTCTTGAACACACAGCGGAGAGGCC 2759
Db 2602 TCTCCTTCTGGATATCTGCTTCAGTGCACAGAACTCTTGAACAAGTAGAGCTGAAGCC 2661
Qy 2760 ATGCCACCTGTTTGAAGAGTCACTGAAAGAGCATCATTTGC---AGAAGAAAATCTCTGTG 2816
Db 2662 GCACCTGTAACTGAGGAGTATTGGAAAGAGAAAGTAATTGCAGAAAGAAACCCCCACG 2721
Qy 2817 CTACCCAGAGCTTACAGAGGGTAAAGATGCCATGACGACATGTCACCACTGAAGTG 2876
Db 2722 GTTACTGAACCTTGCAGAGAACAGAGAGGCCCGGGGCGACACGGTCTGTTAGTAGGGC 2781
Qy 2877 GATTTCACTCAGAACTGTGACAGCCACAGAGACCTCAGAGCTCTCCGTACTGAAGAA 2936
Db 2782 GAATTCACCCCGAAGCTGTGACAGCTGCAGAACTGCAGGGCCATTGGTGCAGAGAA 2841
Qy 2937 GTTACCGAAGCATCGGGGCGGAAAGACAACAGACATGGTGTCCGAGTTTCCGAGCTG 2996
Db 2842 GGAACCGAAGCATCTGCTGCTGAAGAGACCAAGAAATGGTGTGACAGCTCTCCAGTTA 2901
Qy 2997 ACTGACTCCCCAGACACCAAGAGGAGCCACCCAGTTCAGAGGTAGAGAGTGTGTG 3056
Db 2902 ACCGACTCCCCAGACACCAAGAGGAGGCCACTCCGCTGCAGAGGTGGAAGGTGGCGTA 2961
Qy 3057 CTAGATACAGAAAGAGAGGAGCCAGAGCGAGGCCATCTCTCAAGCGTTGCAGACAAG 3116
Db 2962 CTTGACATAGAGAGCAGAGAGGCGGACTCAAGAGTCTCTCAGGAGTGGCAGAAAAA 3021
Qy 3117 GTGAAAGAGAGTCCAGGTGCTGCAACC---CAGACTGTGCAGAGAACGGGGTCAAAA 3173
Db 3022 GTGAAAGAGGATCCCAAGCTGCTGGCACCGGTGGCCAGAAAGTGTGCTTCAAGCTGTG 3081
Qy 3174 GCACTCGAGAGTGTAGAGAGGTAGAGGAGGACTCCGAAGTGTGCTTCGGAAGAAAG 3233
Db 3082 CAGAGAGCAGAGCAGAAAGACCCAGAAAGAGCAGGCTGGAAGCGTCTGGAAGAAAG 3141
Qy 3234 AAGGACGTTATGCCAAAGGACCCGTGCAGGAGCTGGAGCTGAGCATCTTGCACAGGGC 3293
Db 3142 ACCGATGTAGTGTGAAGTAGATGCTCAGGAGGCAAAAACCTGAGCCTTTTACAAAGGG 3201
Qy 3294 TCTGAGACTGGACAGGCTACTCCAGAGAGCCTTG---AAGTTCTTGAAGTCAAGGCAGAT 3350
Db 3202 AAGGTGTTGGGCGAGACCCACCCAGAAAGCTTTGAAAAAGTCTCTCAAGTACAGAGAGC 3261
Qy 3351 GTAGAC-----CATGTGCCAGCTGCCA-----GGTTATCAAG 3383
Db 3262 ATAGAGTCCAGTGAGCTTGTAAACCACTTGTCAAGCCGAAACCTTAGCTGGGGTAAAAATCA 3321
Qy 3384 CTCCAGCAGCTGATGGAACAGGCGGTGGCCCTGAGTCACTCGAAACCTTGCAGACAGT 3443
Db 3322 CAGGAGATGATGGAACAGGCTATCCCCCTGACTCCGGTGGAAACCCCTACAGACAGT 3381
Qy 3444 GAGACAAATGGAGACATCCCTTTAGCAGATTTCAGACATCTGCAGATGGGACACAGCAAGAT 3503
Db 3382 GAGACTGATGGAGACACCCCGTAGCCGACTTTGACGACACCGGACACACCCAGAAAGAC 3441
Qy 3504 GAAACCAATTGACAGCCAGGACAGTAAAGCCACTGAGCTGTGTCAGGAGTCAAGGTCA 3563
Db 3442 GAGATTGTGAAATCCATGAGGAGATGAGTGCATCTGGTACCCAGTCCAGGCGGACACA 3501
Qy 3564 GAGAGAGGCGCTACTGCTCAGAAAGAGGAGCCTTCCACACTACTAATATGTTCCA 3623
Db 3502 GAAGCAGAGGCGAGTTCTCTGCACAGAAAGAGAGCGCCTCCAGACCTTCCAGTTTGTGTTTC 3561

QY 3624 GCCCAGGAGACATGCGGAGAAACGAGAAAGATGTTCTTGAACCTACAGCAAGAG 3683
 Db 3562 CAGGAGAACTAAAGAACATCAAGATGAGAGACCTTAGAGCATATTAAGAG 3621
 QY 3684 CTTACGTGAGCGCGTCCGTTCTGCAAGAAGTGAAGTGGCTCAAGAGGTGAGTT 3743
 Db 3622 GTGTAGTGAAGAACTGTATCACTTCTGCAAGACTGAGGGGACTCAAGAGGTGACAG 3681
 QY 3744 GACTGTGTGATGAGAAAAAGTCAAA-----GAGAAACAGAGGTGTTGTACAC 3794
 Db 3682 TATGTGTATGAGAAAAACCAAGAGTACATTTTTCGAAGGACTTGAAGGGGTCTATAGAC 3741
 QY 3795 TCTGG-----ACCCAACATCAAAAGGCTGTGTATGACATATGACAGTGAAGTATG 3848
 Db 3742 ACAGGATATACAGTCACTAGTCGGGAAAAAGGTCACTGAAGTTGCCCTTAAAGGTGAAGGACA 3801
 QY 3849 GGAGTGGCGCGGCTGTCAAGAAAAAGAGTACTGAAGTCAAGTCTTAG----- 3898
 Db 3802 GAAAGAGTGAATGTAAAAAGATATATGCTCTTGAATCTGCAAGTCAAGCTTAAGTCTCT 3861
 QY 3899 -----CTGAGAGAGGAGAGATGAGAACTGACGTTGAAAAAGAGAAAAAGGAGACAAAG 3953
 Db 3862 CCATCCCCCGGTGAGAGAGAGATGTATGCTCTTGAATCTGCAAGTCAAGCTTAAGTCTCT 3921
 QY 3954 CCAGAGCAATGATGATGAAG---AAGGTAGCAGGAAACAGCCGCTCTGAGCATGAAGA 4010
 Db 3922 CCAAACCATGTGATGAAGAGAGAGAGCTTGAGCAAGAAACAGCTGTATCCGTATCTAAGAG 3981
 QY 4011 ACCTAGGGAAAGCCAGTCTGACACTTGACATGCCAGCTCAGAGAGGGGAGGACACTG 4070
 Db 3982 GTCAATAGCACTCTCTCAGACAGTGAATGTGCCATCATATGATGGGCAAAAGAGATC 4041
 QY 4071 GGAAGCTTTGAGAGAGCCCTTCTCT---CCGAGACCAAGACAAAGCAGTTGATAGAG 4127
 Db 4042 AGCATTTTGAAGAGAAAGCCCTCTCTCCGCTTAGGTCAAGAGAGAGATATGACCAAA 4101
 QY 4128 GTTCAAGTTCAAAAGCTTGAGACAAACAGTCACTCAAAACGCGAAGCTGTGAAAA---- 4183
 Db 4102 ATTCAAGTTCAAGAGCTGTGAGCATCATCTTAACGCGCTGCAAGAGAGGAAAAAG 4161
 QY 4184 --GCTCATGAAACGAGTTGATTTAGAGACAGGTGAAGTCCAGAGTGTAGGTGCA 4241
 Db 4162 GTCTTAGGAGAACTGCAACATTTTGAAGAGGTGAAGCTTGAAGCTCAGAGTGA 4221
 QY 4242 CACTTATTAACAGTGAAGAGTCTCTGCAACGGGTGGCACTGACCTTTGACATGCA 4301
 Db 4222 CATTTAGTTCTGGAAGAGAAATCTCTGAAAAAATGAAGACTTTGCGCTCATCAGAGG 4281
 QY 4302 GAGGACACGGTACCCCTGGGGCTGAGTCTCAGGCGAAGATTCATCCCATCATATGAACT 4361
 Db 4282 GAAAGTGTCTGCTCCCAACAGGGCCGACTGTCAAGCAAAATCGACACAGTGAATGATCT 4341
 QY 4362 CCTGCTCCGGAAGACCCCTACATCTGACCTAACAGAGAAATAAGCCATCCAGAGA 4421
 Db 4342 GCTACTACCAAGAAAGGCTTAAAGTTGACCTGAGAGGAGAGAAACCAATCACTGAAG 4401
 QY 4422 GAGCGATCAGAGAAAGAGACAAAGCAGATGCTGTCTGTATGCTTGAACGCGCAAGAGAT 4481
 Db 4402 TGGAAATCAGATGAATGATGATAGCAGGTGTGCTGCC-----AGAGAGTCAAAAGTGAAT 4455
 QY 4482 ACAGCAATGAAAAAGTCTCTCAAGGCTGAACCTGAGATCTTGAACTTGAAGTAAAGAGC 4541
 Db 4456 GTAGCAATGAGAGGATTTAGAGCTGAAAAATGGGATTTTGGAACTTGAAGACCAAAAGC 4515
 QY 4542 AACAGATTTGTGTGAACGTCATTCAGACAGCCGTTGACCAAGTTGCCAGTAC---AGAA 4598
 Db 4516 AGTAAACTTGTCCAAACATCATCCAGACAGCCGTTGACCAAGTTGTATGATCAGAAAGAA 4575
 QY 4599 ACAGCCCGGAAACTCATGTCTTATGATTCACAGACCCAGGTTG-----C 4642
 Db 4576 ACAGCCACGAAATGTTGAGCTCTGAGTTACAGACCAAGCTCAGGTGATTAAGGTGAGC 4635
 QY 4643 TGCATGACAGCTTTGACAGCAGGAGGCCCAACAGATGCTGACAAAAATGAAGATGCCAA 4702

Db 4636 AGCCAGAGCGCTGGACAGGAAACGAGAAAGAGAGAGAACCTCAGGCTCTGACAG 4695
 QY 4703 GATGAACACCCAGTCCGACGCCA-----GAGAGACCTTGCA 4741
 Db 4696 GATGAACACCAATTAATCTCAGCCAAAGAGAGTCAAGTCAACCGCAGTGGAGCAAGCA 4755
 QY 4742 AGTCTCAACGTTCTGAGAGCATAGGGCTCAGCCCTGGAAATGTGTGCGGCTTGACGTT 4801
 Db 4756 CATTTGATATTTTCCAAAGACATGAGTGAAGCCCTCAGAAAGACATGACTGTTGAGTGA 4815
 QY 4802 GAAAGCCCGGTGTCA----- 4818
 Db 4816 GAAAGTTCACTGTAATATGATCAGACAGCTGGAAGAGGTGCTCCCATCTGAGAGAG 4875
 QY 4819 -----AGTA 4822
 Db 4876 GAGGTGAGCTGGAACAAAGTGTGTGCAAGATAGTGTATGCTTTGACAGAA 4935
 QY 4823 AGCATTGAAGAGCTGCTCCCAACCCAAAG---TCCAAAGAGACATGCTGTGATGCG 4880
 Db 4936 AGAATGAGAAAGTCACTAGTTGAACGAAAGAGATGAAGAAAGGTGATGTTGATGAC 4995
 QY 4881 CCTCAGCTCCAAAGCTTAGCCCAGGCAAGGCGCAGTCTCTGAAAACTTAACCAAGAA 4940
 Db 4996 CTTGAAAAACGAAGTCAAGCCCTGCTGATATGATGCTCTCAGAGAGGCTTAACCAAGAG 5055
 QY 4941 TCCCAACACCAACGACCAAGCTTAACGAGAGGCGATCCCCCAAAAGTTGAGTGC 5000
 Db 5056 TCCCAATCAATCAATGACCAAAACAAAGAGAGAGATGCTCCAGAAATGTGAATTG 5115
 QY 5001 CAGGAAGAGAAATGTCTACCAAGTCAAGTCAAGAGAAAGAGGCCGAGAGAGAGAC 5060
 Db 5116 CAGGAAGAGAAAGTGCACAGTGAATCAGATTAACGATCACACCCAAAGCACAGAGAG 5175
 QY 5061 CTGACAGAGCCAAAGGAGACCTGCGAAGTCTTAAGATGTTAGTTG----- 5107
 Db 5176 TTACGAAACAAAGAGAGAGATCTGCAAGTCAAGATCTTAAGATCATCATG 5235
 QY 5108 -----CTCATTTGATCATCTGTGAAGCCAGATGTGAAGAAACAAAGTCAAGAACAG-- 5157
 Db 5236 CAGTTAAACTATTTGTCTGTTTGAAGCCAGAAATGTGAAGACAAAGTGAAGAAAT 5295
 QY 5158 --AAGCTGCTGTGGAGCCTTGAGACCAAGATTTCAAGAGCCATGAGATCCAGAGAGAG 5215
 Db 5296 GAATGTGCTGTGAGA-CTAAGACAGTATTTTCAAG-ACTTGAAGATTTGAGAGAGAG 5353
 QY 5216 GAGCGTCCAAATGATTTCCACCAGTAGAGCACCCGACAAATTTGAGGCTTCACTGGAG 5275
 Db 5354 GCACATCAACTGATCTCATTTCTAGAGAGC-CCGACCAATCTGAGGCTTCATCAGAGAG 5412
 QY 5276 CTAGAGCCAGCTAACATTTCTCGTTTCAAGCTGCTTGAATTTGCCCTTGATGCCGT 5335
 Db 5413 CTAGAGCCATTTAACATTTCTCTTCCMAAGCCAACTTAACAATTTTCCCTGTAAC-- 5470
 QY 5336 CCGTATATTTCTAACATTTCTCGTTTCAAGAGCTGCTTGAATTTGCCCTTGATGCCGT 5395
 Db 5471 ----- 5470
 QY 5396 CCGTATTTGCGATTTAAGGCTGCGCTTCAACCTGAAC-CAATTTGCGCATACT 5454
 Db 5471 CATATAAATTCGATTTAAGGCTCTAAATTTCTTAACCTGGAACCTGAGATTTGGCAATACT 5530
 QY 5455 AGTTCACTTCTCAAACTGAGACATCTCTTTATGATTTATATGATGATTTATGAG 5514
 Db 5531 AGTTGCGCTTGAAACAGGAGATCATTTCTTAATTTATATGATGATTTTAAAGTAG 5590
 QY 5515 TCTCTCTCTGATCTATTTATATTTTCTTAAGCTTT-----AAG 5557
 Db 5591 ---TCTCTCTGATCTATGATATATTTTCTTAATTTTAAAGAAATGTCAGATATC 5647
 QY 5558 CACATGCTTTTGTATTTATGCAATATATTAACGGGTGTGACGATAGGACGCTTTGAAA 5617

QY 1122 GTTGTGATCTGGAGCATCGTCAAGAGAAAGTGTCTCTTTGGCAACGAAAGTCTT 1181
Db 1341 GTGAGTGGCTCGCAGGGGACCTTTGAAAGAAACCTGCTCCGTGGCAAGAAAGTCTT 1400
QY 1182 GATGAGAAAGTGAAGGCCACCAAGAA---GTTGTGACAGGTCACGTAGACACCGTG 1238
Db 1401 GATGAGAAATGAAAGTCAACAGAAAGGTTGTGGCCGAAGTCCAGTCAAGCCGTG 1460
QY 1239 GAGAAAGCAGAGAGAGAGCAGGAGAGAGAGAGAGCTGAAGGGGGCGTGTGTAGAA 1298
Db 1461 GAGAGAGAAACCGAAGACAA---ACGAGAGTGA 1496
QY 1299 GGAACAGAGAAATCTTCTCCCTGAGAACTGCTGAGCCCGAGAGTCCCCAGAA 1358
Db 1497 GAAACAGAGAGTGTGTCCAGCTGAAGAAATGTGTGAATGAAATGCAAGAACTTCAGAA 1556
QY 1359 GCTGAGCCTGCTGAGGAGCTGATGAAGAGCAGAGATGTGTGTCTGAGGAGACAC 1418
Db 1557 GCCGAACCTGCAAGAGAGCTGAGTGAAGCTCAAGAAAGTGTGTTCGAGAGAGACCT 1616
QY 1419 ACTCAACTGACAGACCTAAGTCTGAAAGAAAGACGCTGCCAAACCCAGAAAGCAT 1478
Db 1617 ACACAGGAGAGTGAAGTCAAGTCTGATGAAGAGTGTGTCCAAACCCCGAAGCGCT 1676
QY 1479 GTCAAGTGAAGTGAAGTGTCTCTCTCAAGAAAGATCAAGGTACAGGAAAGTCTTG 1538
Db 1677 GTGAGTGAAGTGAAGTGTCTCATCAAGAGAAATGAAGGTGCAAGGAAGTCTACTA 1736
QY 1539 AAGAAACCTTCATGATGCTAGGCTTAAAGAAAGTCTGAGGAAAGAGAGAGAGAA 1598
Db 1737 AAGAAAGCTTTTACAGACATGAGCTTAAAGAAAGCTTTGAAAGAAACAGAAAGGAAA 1796
QY 1599 CGAGAGGTGGGGAGAGAGAGAGCTGAGAAATACAAACATTCACAGCGAATCCCA 1658
Db 1797 AGAGAA---GGAAGAGAGAGAGAAATCAGGGAGACACATCAGGTTCCAGCGAATCTCG 1853
QY 1659 GAGAGTGTGATGAGCAGAGAGGAGAGCTCTGCTGCTGCTCCCGAGAGAGCTTGAAG 1718
Db 1854 GACAGCCAGAGAGAGCAAAAGGGCGAGAGCTCTGCTCATCCCTGAAGAGGCCAGAGAG 1913
QY 1719 ACCAGTGTGAGAAAGAGGGCGCTGGAAGACCCCAAGATGGGAAAGCTGAGAGAA 1778
Db 1914 ATCAGTGTGAGAAAGGGCTTACCCAGAGTGCAGAGAGATGGGAAAGCTGAAGAGAA 1973
QY 1779 ACTACTTCGATGAGAGAAAGAGAGAGGATCACTCCCTGGGCACTCTTCAAAAAG 1838
Db 1974 GCTACTTCGATGAGAGAAAGAGAGAGGATCACTCCCTGGGCACTCTTCAAAAAG 2033
QY 1839 ATGTGACACCCCAAGAAACGGGTCCGAAGACCTTCTGAGTGAAGAGAGAGAGCTG 1898
Db 2034 ATGTGACGCGCCCAAGAAACGTTTGAAGCGCTTCGAAAGTGAAGAAATAGATGCTG 2093
QY 1899 GAGAAAGTCAAGAGGCGCACTGTGCTCCACTGATGACAGGTGCAAGAAATGCAAGT 1958
Db 2094 GACAAAGTCAAGAGGCTTACTTGTCTCACCGAGAGACAGCTCTGAAATGCAAGAA 2153
QY 1959 GAAGTCAAAACCTGTGTGAGAGAAACAAAGCCAGAGAAACCAAGCTTGAAGTGAATCT 2018
Db 2154 GAAATGAAGAGAGGCTGAGAGAGCAAAAGCCGAAAGAACAAAGGTGAGTACC 2213
QY 2019 TCAGTGTCTTGGAGAGCACTGATTTGTGTGATCATCCAGAGAGAGAGAGCA 2078
Db 2214 TCAGTATCTTGGAGAGCTTTAATTTGTGTGATCATCAAGAAAGAGCAAGGAGAAAG 2273
QY 2079 TCTCTTTCAGATGAGAGAGAGGCGCAAGACATGAGAGGAGAGAGTCAAGACAGAG 2138
Db 2274 TCTCTTTCAGATGAGAGAGGAGAGCAAAAGCAATGGAGAGAGACCAAGAAAGCTGAT 2333
QY 2139 GAGGCGCAAGAGCAAG 2198
Db 2334 GAGGCGCAAGAGCAAG 2393

QY 2199 GACCAAGCGAAGGAAGTTCCTCAACCCAGCAGCGGAAAGCCCTTCCGAAGGGGAAGT 2258
Db 2394 GATCAGAGGAGAGAGATCTCTCCCGAGACAGCTGAAAGCCCTACCAAGGGAGAGGC 2453
QY 2259 GTCTCACTTGGAGAGTATTTAAAGATTAGTCACTCCAGAAAGAAATCCAGTCAAA 2318
Db 2454 GTTTCACCTGGAGAGTCAATTTAAAGTTAGTCAAGCAAGAAAGAAATCAAGTCCAG 2513
QY 2319 CTGGAAGAAAGCCGAAGAC-----TCTAGTGAAGCACTGTTCCAGTGAAG 2366
Db 2514 CTGGAAGAAAGAGCAAGACTCCATAGCTGAGTCTGTGTAGAAACATTCACCTCAGAC 2573
QY 2367 ATGGAACGAGTGAAGAGATCTGGGTTTCCATTAAAGAAATCATCCCGAGCGAG 2426
Db 2574 ACTGAACCCGGAAGAAAGTCTGGGTCTCAATCAAGAAAGTTATTCCTGAGAGAG 2633
QY 2427 AAGAAAGGCGAGACGGGAGCAAGAAACAGACCTGTGAAAGACTGAGGCGCAGTGAG 2486
Db 2634 AAGAAAGGCGAGATGGGAAACAAAGAACCCCTGTGAAAGCAGAGGCGCAAGAGG 2693
QY 2487 ATAAATGAGACGACCTTAATGTCCAGCCGTCGCTCTGTGTGATATGCAAGT 2546
Db 2694 GCCAAGAGATGACTGTGATGTCCGCGGTGCTCTGTGTGATGATGATGATG 2753
QY 2547 GAGAGGAGAAAGATG-----AAGCCAGGGAAATACGAGCTGCGCCAGCTGCTG 2597
Db 2754 GAAAGGAGAAATGAGAGCAGACAGAGCCCAAAAGCGTAGAGAGCCGAGACAGAG 2813
QY 2598 GGGGCTGTGTACTGTCTGAGAGAGCTGTAAGACTGTGTCAACTGTGATGTGCA 2657
Db 2814 GCAGCAGATGAGTGTCAAGAGAGCTCAGGAGAGTCAAGTGTATATGATGAGAGCT 2873
QY 2658 GTCAATGATGAGCAGAGGAGTCAACAGTGTGAAGAGCGTCTCTGTGTGATATCC 2717
Db 2874 GTGCTGAGGAGAGAGAGGCGACTTACATTAAGAAAGGTCTCTTCTGTGATATCT 2933
QY 2718 GCTTCCGTAAGAAACCTCTTGAACACACAGCGGAGAGCATGCCACTGTGTAAGAG 2777
Db 2934 GCTTCAAGTGAAGAACTCTTGAACAGTGAAGCTGAAGCGCATGTTAATGAGAG 2993
QY 2778 GTCACTGAAAGACATCATTTG---AGAGAAACCTCTGTGTCAACCAAGCTTACA 2834
Db 2994 GTATTGAAAGAAAGTATTTGCAAGAAAGAACCCCAAGTTACGAACCTGTGCA 3053
QY 2835 GAGGTAAGATGCCATGACGACATGCTACACATGATGATTTCACTCAAGAGCT 2894
Db 3054 GAGAAACAGAGAGGCGCGGAGCAGACCGTCTGTAGTGGCGGAGATTGACCCCGAAGCT 3113
QY 2895 GTGACAGCAGAGAGACTCAAGAGCTCTCGTACTGAAGAGTTACGAAGCATGCGGG 2954
Db 3114 GTGACAGCTGAGAAATGACAGGAGCATTTGGGTTCGAAGAGAACGAAGCATGTGCT 3173
QY 2955 GCCAAGAGACCAAGACATGTTGCGCAGTTCCTCCAGCTGACCTGCCAGACAC 3014
Db 3174 GCTGAAGAGACCAAGAAATGTTGTGACAGTCTCCAGTTAACGACTCCCAAGACAC 3233
QY 3015 AAGAGAGACCAACCCCAAGTTCAGAGGTTAGAGAGTGTGCTGATACGAAGAGAG 3074
Db 3234 ACAGAGAGGCGCACTCCGTTGACAGAGGTGAAGGTGCTGATCATGATGAAGAGCA 3293
QY 3075 GAGGCGCAGAGAGGCGCATCTCAAGCGTGTGACAGCAAGGTGAAGAGAGAGTCCAG 3134
Db 3294 GAGAGGCGGACTCAAGAGGTTCTCAAGCAATGAGAGAAAGTGAAGAGATCCAG 3353
QY 3135 GTGCTGCAAC---CAGACTGTGAGAGAAAGGGGTCAAAAGCATGGAAGAGTTGAG 3191
Db 3354 CTGCTGCGACAGGAGGCGCAGAAAGTGTGCTTCAAGCTGTGAGAGAGAGAGCA 3413
QY 3192 GAGGTGAAGAGACTCGAAGTGTGCTTCCGAGAAAGAGAGAGAGAGAGAGAGAG 3251
Db 3414 AGACCAAG 3473
QY 3252 GAGCCCGTGAAGCT 3311

Db 3474 GTAGATGCTCAGGAGGCAAAACTGAGCCTTTTACAAAGGGAAGGTGGTGGGCGACACC 3533
Qy 3312 ACTCCAGAGAGCCTTG---AAGTTTCTGAAGTCAAGGAGAGTGTAGAC-----CAT 3359
Db 3534 ACCCCAGAAAGCTTTGAAAGAGCTCCTCAAGTACAGAGAGCATAGAGTCCAGTGAGCTT 3593
Qy 3360 GTCCGCAAGTGCCA-----GGTTATCAAGCTCCAGAGAGTGTATGGAA 3401
Db 3594 GTAAACCACTTTGTCAAGCGCAAAACCTTGTAGTGGGGTAAAAATCACAGGAGATGGTATGGAA 3653
Qy 3402 CAGCCGTGGCCCTGAGTGCATCCGAACCTTGACAGACAGTGAAGAAATGGGAAGCACT 3461
Db 3554 CAGGCTATCCCCCTGACTCGGTGGAAACCCCTACAGACAGTGAAGTGTATGGGAAGCACC 3713
Qy 3462 CCCTTAGCAGATTTCAGACACTGCGAGATGGGACACAGCAAGATGAACCATTTGACAGCCAG 3521
Db 3714 CCGGTAGCCGACTTTGAGCGCACAGGACAAACCCAGAAAGACGAGATTGTGGAATCCAT 3773
Qy 3522 GACAGTAAAGCCACTGAGCTGTTCAGGAGTCAAGGTTCACAGAAGAGAGCGGCTACT 3581
Db 3774 GAGGAGAAATGA---GGTGCATCTCGTATCCAGTCAAGGGGCACAGAAGCAGAGCGAGTTCCCT 3830
Qy 3582 GCTCAGAAAGAGGAGCCTTCGACACTACCTTAATAATGTTCCAGCCAGGAGAACATGGG 3641
Db 3831 GCACAGAAAGAGAGGCCCTTCAGACACTTCAGAGCTTGTGTTTCCAGGAAGAAACTTAAAGAA 3890
Qy 3642 GAAGAACCCAGGAAGAGATGTTCTTGAACCTACACAGCAAGAGCTTACTGTGTCAGCCGTG 3701
Db 3891 CAATCAAGATGGAGACACTCTAGAGCATACAGATAAGAGGTGTCTAGTGGAACCTGTA 3950
Qy 3702 CCGGTTCTGGCAAGACTGAGGTGGGTCAAGAGGGTGA-----GGTTGACTGGTTG 3752
Db 3951 TCCATTCTGTCAAGACTGAGGGGACTCAAGAGCTGACAGTATGCTGTGATGAGAAAC 4010
Qy 3753 GATGGAGAAAGTCAAGAGAACAGAGGTGTTGTACACTCTGG-----ACCCAAC 3806
Db 4011 AAGACGTACCATTTTTCGAAGGACTTGAGGGGTCTATAGACACAGGCATTAACAGTCAGT 4070
Qy 3807 AGTCAAAAGGCTGTGATGTGACATATGACAGTGAAGTGTGGAGTGGCGGGTGTGAC 3866
Db 4071 CGGGAAGGTCACTGAAGTTGCCCTTAAAGGTGAAGGACAGAAAGCTGAATGTAAA 4130
Qy 3867 GAAAAGGAGAGTACTGAAGTGCAGAG-----TCTTAGCCTGGAGGAGGA 3911
Db 4131 AAGGATGATGCTCTTGAACCTGCAGAGTCACGCTAAGTCTCTCTCATCCCCCGTGGAGAGA 4190
Qy 3912 GAGATGGAAGTACGTTGAAAGGAGAAAGGAGAGACAAGCCAGAGCAAGTGAAGTGAA 3971
Db 4191 GAGATGGTAGTTCAAGTCGAAGAGGGAGAAACAGAAAGCAGAGCCCAACCCATGTGAATGAA 4250
Qy 3972 GAAGG---TGAGCAGGAACACAGCCGCTCTGAGCATGAAGGAACCTACGGGAAGCCAGTC 4028
Db 4251 GAGNAGCTTGAGCAGCAACAGCTGTTACCGTATCTGAAGAGGTCACTAAGCAGCTCCTC 4310
Qy 4029 CTGACACTTCACATGCCAGCTCAGAGAGGGGAGAGCACTGGGAAGCCTTTGGAGGAAGC 4088
Db 4311 CAGACAGTGAATGTGCCCATCATAGATGGGGCAAGGAAGTCAGCAGTTTGGGAAGGAAGC 4370
Qy 4089 CTTTCTC---TCCAGACCAAGACAAAGCAGGTGTGATAGAGTTCAAGTTTCAAGCCTG 4145
Db 4371 CTTCTCTCCCTGCTTAGTCAAGAGAGGCGAGTATGCACCAAAATTTCAAGTTTCAGAGCTCT 4430
Qy 4146 GACACACAGTCACTCAAAACAGCAGAGAGCTGTGAAAA-----GGTCATAGAAACCGTT 4199
Db 4431 GAGGCATCTTCTACTTAAACAGCGGCTGCAGAGAGGAAAGGTCTTAGGAGAACTGCC 4490
Qy 4200 GTGATTTTCAGAGACAGGTGAAGTCCAGAGTGTGTAGGTGCACACTTATTAACAGCTGAG 4259
Db 4491 AACATTTTAGAAAACAGGTGAAAACGTTGAGCCTGCAGGTGCACATTTAGTTCTGGAAGAG 4350
Qy 4260 AAGTCTCTGAAACGGGTGGCCACTGAGCTCTTCAGCATGTCAGAGGACAGCGTACCCCTG 4319

Db 4551 AANTCCTCTGAAAAAATGAAGACTTTTCCGCTCATCCAGGGGAAGATGCTGTGCCACACA 4610
Qy 4320 GGGCCTGAGTCTCAGGCAGAAATCCATCCCAATCATAGTAACTCTCTGCTCTGAAAGCACC 4379
Db 4611 GGGCCCGACTGTTCAGGCAAAATCGACACAGTGTAGTATCTGCTACTACCAAGAAAGGC 4670
Qy 4380 CTACATCTGACCTACAGGAGAAATAAGCGCATCCAGAGAGAGCGATCAGAGGAAGAG 4439
Db 4671 TTAAGTTTCCGACCTGGAAGGAGAGAAAAACCATCACTGAAGTGGAGTTCAGATGAAGTC 4730
Qy 4440 GACAAGCCAGATCCTGCTCTGATGCTGACGCAAGGAGAGTACAGCAATTCGAAAAAGTC 4499
Db 4731 GATGAGCAGGTTCCTTGCC-----AGGAGTCAAGTGTAGTGTAGCAATTG---AGAT 4781
Qy 4500 CTCGAAGCTGAACCTGAGATCTCGGAACCTTGAGAGTAAGAGCAACAGATTTGTCTGAAC 4559
Db 4782 TTAGAGCCTGAAAAATGGGATTTTGGAACTTGAGACCAAAAGCAGTAACTTTGTCCAAAAC 4841
Qy 4560 GTCAATTCAGACAGCCGTTGACCAAGTTGCGACGTAC---AGAAAAGCCCGCCGAAATCAT 4616
Db 4842 ATCATCCAGACAGCCGTTGACCAAGTTTGTACGTACAGAAAGAAACAGCCACCGAAATGTTG 4901
Qy 4617 GCTTATGATTTTCAGACCCAGGTTT-----CTGCATGCGAGGCTTGACAG 4660
Db 4902 ACGTCTGAGTTTACAGACACAAGCTCAGTGTATAAAGCTGACAGCCAGGACGCTGGACAG 4961
Qy 4661 CAGGGAGCCCAACAGATGCTGGGCAAAAAATGAAAGATGCCAAGATGAAACACCCAGTGCC 4720
Db 4962 GAAACGGAAAGAGAGGAGGAGAACCTCAGGCCCTCTGCACAGGATGAAACACCAATTACT 5021
Qy 4721 GCAGCCCCA-----GAGAGGACTTGCAGTCTGCAAGTCTCTGACCGTTCTGGA 4759
Db 5022 TCAGCCAAAGAGGAGTCAAGTCAACCGCAGTGGGCAAGCAATCTGTATATTTCCTCAA 5081
Qy 4760 GGCATGGGCTCAGCCTCGGAAATGCTTGGCGCTTGCAGTTGAAAGCCGCGGTGTCAA- 4818
Db 5082 GACATGAGTGAAGCCTCAGAAAAAGACCATGACTGTTGAGGTAGAAGGTTCCACTGTAAAT 5141
Qy 4819 ----- 4818
Db 5142 GATCAGCAGCTGGAAGAGGTGCTCTCCCATCTGAGGAAGAGGAGGTGGAGTGGAAACA 5201
Qy 4819 -----AGTAAGCATTTGAGAAGTGCCT 4840
Db 5202 AAGTCTGTGCGAGAAGATGATGCTCATGCTTGTAGCAAGAAAGATAGAGAAGTCACTA 5261
Qy 4841 CCTCAACCCAAAGA--TCCAAAGGAGCATGCTGTGATGGCCCTCAGCTCCAAAGCTTA 4898
Db 5262 GTTGAACCCGAAAGAGATGAAAAAGGTGATGATGTTGATGACCTGAAAAACAGAACTCA 5321
Qy 4899 GCCCAGCAGAGGCCAGTGCCTCTGGAAACCTTAACCAAGAAATCCCAGACACCAACGGA 4958
Db 5322 GCCTTGCTGATACTGATGCTCAGAGGCTTAAACCAAGAGTCCCAGATACAAATGGA 5381
Qy 4959 CCAAAGCTAACCGAGGAGGCGATCCCAAAAGTTGAGGTCCAGGAAGAAATGTCT 5018
Db 5382 CCAAAAACAAAAGAGAGAGGAGGATGCCAGGAAGTAGAATTGCAGGAAGGAAAGTGCAC 5441
Qy 5019 ACCAAGTCACTAAAGAGAAACAAGGCCCGGAGAGAGAGACCTGCAGGAGCCAAAGGGA 5078
Db 5442 AGTGAATTCAGATAAGCGATCACACCCCAAGCAAGGAGGAGTTACAGAAACCAAGAGAGA 5501
Qy 5079 GACCTGCGAGAACTCCTAAGATGTTAGTTG-----CTCATTTGTAC 5117
Db 5502 GAATCTGCAAGTTCAGAACTTACAGAACTTAAAAACATCATGCAAGTTAAACTCATTTGCT 5561
Qy 5118 ATCTGTAAGACCAAGAATGTGAAAAACAAAGTCACAGAAACA-----GATGCTCTGCTGGGAC 5173
Db 5562 GTTTGGAGACCAAGATGTGAAGCAAGTAGTAGAAGAAATGAATGCTGCTGCTGAGA- 5620
Qy 5174 CTTGAGACCAAGATTTTCAGAGCCCATGAGATCCAGAGAGCAGGCGGTCCAAATGATTTCC 5233
Db 5621 CTGAAGACCAAGTATTTTCAGA-ACCTTTGAGAAATTCGAGAGCAGGACATCAACTGATCTCA 5679

QY 357 AACTCCACAGCTGTTGAAGATATCACAAAGAGTGGCGAGGAGGACATCAGAAATAATT 416
Db 258 AAGTCAGCGGTGTTTCAGACATCACAGATGATGGCGAGGAGGACACCCGAAATAATC 317
QY 417 GAACAGATCCCTGCTTCAGAAAAAATAATGTGGAAGAAATGGTACAGCCTGCTGAGTCCAG 476
Db 318 GAACAGATTCCTTCTTCAGAAAGCAATTTAGAGAGCTAACACACCCACTGAGTCCAG 377
QY 477 GCTAATGATGTTGGCTTCAAGAAAGTATTTAAATTTGTTGGTTTTAAATTCACGGTGAAG 536
Db 378 GCTAATGATATTGATTTAAGAGGTTTAAAGTTTGTGGCTTTAAATTCACCTGTGAAA 437
QY 537 AAGGATAAAATGAAGAGTCAGATACCTGCTCAACTACTCTGCTCAAGAGAGATGAAGGC 596
Db 438 AAGGATAAGACAGAGAAGCCTGACACTGTCCAGTACTCTGCTGAAGAAAGATGAAGGG 497
QY 597 GAAGGGCAGAGAGCCTCTGTCGAGCTGGAGACCAACAGAGGCCAGCTGTGGAGACTGCC 656
Db 498 GAGGGAGCACA-----GGGGCTGGCGACCAAGAGNCCCAGCCTTGGGGCT--- 545
QY 657 GTCGAGAGTCAGATCCAAAGAAAGTGAGCTGAAGCAATCCACAGAGAAAGCAAGAGGC 716
Db 546 ---GGAGAAGCAGCATCCAAAGAAAGCGAAACCCAAAATACTACAGAGAAAACCCGAAGAG 602
QY 717 ACCCTGAAGCAAGACAGAGCAGCACAGAAATCCCCCTTCAAGCCGATCTGATCAAGCG 776
Db 603 ACCCTGAAGCGTGAGCAAGGCCAGCGAATAATTTCTCCCCAGCCGAATCTGGCCAA--- 659
QY 777 GCTGAGGAAGAACCAAGATGAAGAGAAAGAAACAAAGAGAAAGGCCACCAAGTCC 836
Db 660 GCAGTGGAGGATTCAAAGAGGAGGAGAGAGAAACAGAAAGAAACCTAGCAAGTCT 719
QY 837 CCAGAAATCCCGAGCAGCCAGTCAACAGTGAGCAACATCTTCTTTCAAGAAATCTTTC 896
Db 720 GCAGAAATCTCCGACTAGTCCGCTGACCAGTGAACAGGATCAACCTTCAAAAAATCTTC 779
QY 897 ACTCAGGTTGGCGCGCTGGCGCAAGACACAGCTTCAAGAAATCAAAAGAGGATGAT 956
Db 780 ACTCAGGTTGGCGCGCTGGCGCAAAAGACCAAGTTTCAGGAAGCCGAAAGGAGGATGAA 839
QY 957 CTGGAATCTCGAGAGAGAAAGAGCAAGAGGCAAGAAAGTGAAGAGGAGAAAG 1016
Db 840 GTGGAGCTTCAGAGAAAGAAAGGAAACAGAGCCAGAAAGATGAGACAGAAAGAGAC 899
QY 1017 GAAAGACAGAGCCAGCTCGGAGGA-----GAGGAG 1049
Db 900 GGAAGGCAGAGGTTGCTTCGAGAAACTGACCGCCTCCGAGCAAGCCCACACAGAG 959
QY 1050 CCGCAGAGACACAGACCAAGCCAGGTTGTGACGACACTACAGAAAGTGGAGTGCCT 1109
Db 960 CCGCAGAAAGTGCCACAGAGCCCGGTTATCAGCTGAATATGAGAAAGTTGAGCTGCC 1019
QY 1110 TTGGAAGACAGGTTGTGACCTGGAGGACCTGTCAGAGAGAGTGTGCTCTTTGGCA 1169
Db 1020 TCAGAGAGCAAGTCAAGTGTGCGGGACCTTCTGAAGAGAAACCTGCTCGCTTGGCG 1079
QY 1170 ACGGAAGTGTGATGAGAAAGTGAAGCCCAAGAA---GTTGTTGAGAGGTTCCAC 1226
Db 1080 ACAGAAGTGTGATGAGAAATAGAGTCCACCAAGAGAGGTTGTGGCGAAGTCCAC 1139
QY 1227 GTGAGCACCGTGGAGAGACAGAGGAGGAGGAGGAGGAGAGGCTGAAGGGGGC 1286
Db 1140 GTCAGCACCGTGGAGGAGAGAACCGAAGAGCAGAAA----- 1175
QY 1287 GTGTGCTGAGAGAAACAGAGGAAATCTTCCCTCTGAGAACTGGCTGAGCCGAGAG 1346
Db 1176 ACGAGGTGAAGAAACAGCAGGGTCTGTGCCAGCTGAAGAAATTTGTTGAATGGATGCA 1235
QY 1347 GTCCCCCAGGAAGCTGAGCCTGTGAGGAGCTGATGAAGAGCAGAGAGATGTGTCTCT 1406
Db 1236 GAACCTCAGGAGCTGAACCTGCCAGGAGCTGTGTAAGCTCAAGAAACGTGTGTTCC 1295
QY 1407 GGAGGAGACCACACTCAACTGACAGACCTTAAGTCTCTGAAGAGAGAGCGCTGCCAAACAC 1466

Db 1296 GGAGAGACCCCTTACACAGGAGGCTGACCTCAGTCCTGATGAGAAGGTGCTGTCAAACCC 1355
QY 1467 CCAGAAGGCATTTCTCAGTGAAGTGGAGATGCTCTCTCTCAGGAAAGAAATCAAGAGTACAG 1526
Db 1356 CCCGAAGGCGTTGTGAGTGAAGTGGAAATGCTGTCTCATCAGAGAGAGATGAAGGTGCAG 1415
QY 1527 GGAAGTCCCTTGAAGAAACTTTCTCAGTGAAGTCTAGGCTTAAAGAGCTGTCTGGAAAGAG 1586
Db 1416 GGAAGTCCCAATAAGAAAGCTTTTACCAAGCACTGGCTTAAAAAAGCTTTCTGGAAGAAA 1475
QY 1587 CAGAAGGGGAAACAGAGAGTGGGGAGAGAGAGCCTGGAAGATACCAACATTCAC 1646
Db 1476 CAGAAAGGGGAAAGAGGA---GGAGGAGACAGGAATCAGGGAGGACACTCAGGTTCCA 1532
QY 1647 ACCGAATCCCCAGAGAGTGTGATGAGCAGAGGGAGAGAGCTCTGCTGCTCCCCCGAG 1706
Db 1533 GCCGATTTCTCGGACAGCCAGGAGGAGCAAAAGGGCGAGAGCTCTGCCTCATCCCCTGAG 1592
QY 1707 GAGCCTGAGGAGACCAAGTGTCTGGAGAAAGGGCCGCTGGAAAGCACCCAGGATGGGAA 1766
Db 1593 GAGCCCGAGGAGATCAAGTGTCTGGAAAGGGCTTAGCCGAGGTGAGCAGAGATGGGAA 1652
QY 1767 GCTGAGAGAGGAACTACTTCCGATGAGAGAGAAAGAGAGAGGATCACTCCCTGGCA 1826
Db 1653 GCTGAAGAGGAGGACTACTTCCGATGAGAGAGAAAGAGAGGATGTCACTCCCTGGGCA 1712
QY 1827 TCCTTCAAAAAGATGTGACACCAAGAAACGGGTCCGAAAGACTTCTGAGAGTGACAAG 1886
Db 1713 TCATTCAAAAGATGTGAGCGCCCAAGAGCGTGTAGACGGCTTCGGAAGTGATAA 1772
QY 1887 GAGGAAGAGCTGGAGAGGTCAAGAGCGCACCTTGTCTTCACTGATAGCACAGTGTCA 1946
Db 1773 GAAGATGAGTGGACAAGGTCAAGAGCGCTACCTGTCTTCCACCGAGAGCACAGCCTCT 1832
QY 1947 GAATGCAAGATGAAGTCAAACTGTTGTGAGAGCAAAAGCCAGAGAACCAAGAGCGT 2006
Db 1833 GAAATCAAGAAAGAAATGAAGGGAGCGTGGAAAGAGCCAAAGCCGGAAGAACCAAGAGCG 1892
QY 2007 AGGTGGATACTTTCAGTGTCTTGGGAAGCACTGATTTGTGTCGATCATCAAGAGAGAGA 2066
Db 1893 AAGTGGATACCTCAGTATCTTGGGAAGCTTTAATTTGTGTGGATCATCAAGAGAAAGA 1952
QY 2067 GCAAGAAAGCATCTCTTCAGATGATGAAGAGGCGCAAGGACACTGGGAGGGGACAGT 2126
Db 1953 GCAAGGAGAGGTCCTCTCTGATGAGGAAGGGGACCAAAAGCAATGGAGAGAGCCAC 2012
QY 2127 CACAGAGCAGAGGAGGCGACAAAGAACAAAGAGCCGGAAACAGACGCTGTCTTCCAGC 2186
Db 2013 CAGAAAGCTGATGAGGCGGAAAGAACAAAGAGACGCGGACAGACGGGATCTTTGCTGCT 2072
QY 2187 ACCCAGGAGCAGGACCAAGCGCAAGGAAGTTCTCTACCCGAGCCAGGGAGCCCTTCC 2246
Db 2073 TCCAAAGAACATGATCAGGGCAGGGAAGTTCTCCCGGAGCAAGTGAAGCCCTTACC 2132
QY 2247 GAAGGGGAGGTTCTCTCACTTGGGAGTCAATTTAAAGATTTAGTCACTCCAGAAAAAAA 2306
Db 2133 GAAGGGGAGGCGTTTCCACTCGGAGTCAATTTAAAGGTTAGTCAACGCCCAAGAAAAA 2192
QY 2307 TCCAGTCAAACTGGAAGAGAAAGCCGGAAGAC-----TCTAGTGTAGAGCAG 2354
Db 2193 TCAAAGTCCAAGCTGAAGAGAGAAAGCAAGACTCCATAGTGGGTCTGGTGTAGAACAT 2252
QY 2355 TTGTCCACTGAGATCCAAACCGAGTAGAGAGATCTTGGGTTTCCATTAGAAATTCATC 2414
Db 2253 TCCACTCCAGACACTGAACCCGGTAAAGAGAAATCTCGGTCTCAATCAGAAAGTTATT 2312
QY 2415 CCCGAGCGCGGAGAAAAAGGGCAGACGGGAGCAAGAAACAAAGCCACTGTGGAGAGACTCA 2474
Db 2313 CTTGGACGAGGAGGAGAAAGGGCCAGATGGGAACAGAGAACAGCCCTCTGTGAGAGCGCA 2372
QY 2475 GGCCCAAGTGGAGATAAATGAGGACGACCCCTAATGTGCCAGCGCTGTGCTGTCTGTAG 2534

Db 2373 GGGCCAAAGGGGCCAAGAAATGACTGTATGTCGCCGCCGCTGCTCCTGTCTAG 2432
 QY 2535 TATATGCACTGAGAGAGGAGAAATGG-----AAGCCAGGGGAAATCGAGCTG 2585
 Db 2433 TATGATGCTGTAGAAAGGAGAAATGGAGGCGACAGCAAGCCCAAAAAAGCCAGAGCAG 2492
 QY 2586 CCCCAGCTGCTGGGGGCTGTATAGTCGCCAGGAGCTCAGTAAAGCTCTGTGTCACACT 2645
 Db 2493 CCGGAGCAAGAGGCGACCACTGAGGTGTCCAAAGAGCTCAGGAGATCAGGTTCAATATG 2552
 QY 2646 GTGAGTGTGCGAGTATTGATGGACCAAGGCGAGTCACCAAGTGTGGAAGAGCGGTCTCT 2705
 Db 2553 ATGGAGCAGAGCTGTGCTGTGAGGAGAGAGGCGACAGCTATTATGAAAGAGTCTCCT 2612
 QY 2706 TCGTGGATATCCGCTTCCGTAACAGAACTCTTGAACACACAGCGGGAGAACCTATGCCA 2765
 Db 2613 TCTTGGATATCTGCTTCACTGACAAACCTCTTGAACAAGTGAAGCTGMAACCCACATG 2672
 QY 2766 CCTGTGMAAGGAGTCACTGMAAAAGACATCATTTGC---AGAAAGAACTCTGTGTCAC 2822
 Db 2673 TTAACTGAGAGGTATTTGAAAAGAGAAATTTGCAAGAAAGAACCCCAACGGTTACT 2732
 QY 2823 CAGACGTTACCAAGAGGTAAAGATGCCCATGACGATGTGACCAAGTGAATTTTC 2882
 Db 2733 GAACCTCTGCCAGAAACAGAGAGGCCCCGGGCGACACGCTGTTAGTGAAGCGGAATTG 2792
 QY 2883 ACCTGAGAACTGTGACAGCCACAGAGACCTCAGAGGCTCTCCGTACTGAAGAATTACC 2942
 Db 2793 ACCCCGAAGCTGTGACAGCTCAGAAACTGCAAGGCGCATTTGGGTGCCAAGAAAGAAC 2852
 QY 2943 GAAGCATCGGGGCGGAGAAAGACCAAGACATGTGTCCGCAAGTTTCCAGCTGACTGAC 3002
 Db 2853 GAAGCATCTGCTGTGAAGAACCAAGAAATGTGTGACAGCTCTCCAGTTAACCGAC 2912
 QY 3003 TCCCAAGACCAAGAGAACCAACCCCAAGTTCAAGAGTGAAGTGTGTGATAGT 3062
 Db 2913 TCCCAAGACCAAGAGAGGCGCACTCCGCTGCAAGAGTGAAGTGTGTGATAGT 2972
 QY 3063 ACAGAAAGAGAGAGGCGCAGACGAGGCACTCTCAAGCCGTTGCAAGAGTGA 3122
 Db 2973 ATAGAAGAGCAAGAGAGGCGCACTCAAGAGTCTCTCAAGGAGTGAAGAAAGTGA 3032
 QY 3123 GAGGAGTCCCAAGTGTCTTCAAC---CAGACTGTGCAAGAAAGGAGTCAAGACTG 3179
 Db 3033 GAGGAATCCCAAGTGTCTTGAAGAGGCGGAGGAGGAGTGTCTTCAAGCTGTGCA 3092
 QY 3180 GAGGAAGTGTGAGAGTGAAGAGGACTCCGAAGTGTGCTTCCGAGAAAGAGAGAC 3239
 Db 3093 CGAGAGGCGAAAGACCAAGAAAGAGAGGCTGAAGCTCGGGTCTGAAGAAAGAGCGAT 3152
 QY 3240 GTTATGCCGAAAGGAGCCCGTGCAGAGAGTGAAGTGAAGTGTGACATCTTGCACAGG 3299
 Db 3153 GTAGTGTGAAGATGATGCTCAGAGGCGAAAGCTGAAGCTTTTCAAGAGGAGG 3212
 QY 3300 ACTGCAAGGCTACTTCAAGAGGCTTGA---GTTCCGAAAGTCAAGCAAGTGAAGAC 3356
 Db 3213 GTGGGCGACAGCCCAAGAAAGCTTGAAGAGCTTCAAGTCAAGAGAGCATAGAG 3272
 QY 3357 -----CATGTGCCACGTGCCA-----GGTTATCAAGCTCCAG 3389
 Db 3273 TCCAGTGAAGCTTGAAGCACTTGTCAAGCCGAAGCTTGAAGTGGGTTAAATCAAGAG 3332
 QY 3390 CAGCTGATGAAGAGGCGCGTGGCCCTGAGTCAATCCGAACCTTGAACAGAGTGAAGAC 3449
 Db 3333 ATGTGATGAAGAGGCTATCCCTCTGACTGTGTGAAGAAACCCCTACAGAGTGAAGCT 3392
 QY 3450 AATGGAAGCACTCCCTTAGCAATTCAGACACTGCAAGTGGGACACAGCAAGTGAAGCC 3509
 Db 3393 GATGGAAGCAACCCCTGATCCGACTTTGAGCGCACAGGACCAACCCAGAAAGCGAGATT 3452
 QY 3510 ATTGACAGCAGGACAGTAAAGCACTGCAAGCTGTCAAGGAGTCAAGGTCACAGAGAA 3569
 Db 3453 GTGGAATCACTAGAGGAATGAGGTGCTGTGTTACCAAGTCAAGGGGCGACAGAAAGCA 3512

QY 3570 GAGCGGCTACTGCTCAAGAAAGAGAGGCTTTGACACTACCTAATTAAGTTCCAGCCAG 3629
 Db 3513 GAGCGAGTCTCTGCAAGAAAGAGAGGCTTCAGACCTTCAAGTTTGTGTTCCAGAA 3572
 QY 3630 GAAGAATCATGGGAGAAACCAAGAAAGATGTTCTTGAACCTTACACAGCAAGACTTACT 3689
 Db 3573 GAACATAAGAACAAATCAAGATGGAAGACACTTAAAGCATTAAGATTAAGAGGTGCA 3632
 QY 3690 GCTGACCGTGGCCGTTCTGGCAAGACTGAGTGGGTCAAGAGGTTGAGTTGACTGG 3749
 Db 3633 GTGGAACCTGTATCATCTGTCTCAAGACTGAGGGGACTCAAGGCGTGCACAGTATGCT 3692
 QY 3750 TTGATGGAAGAAAGTCAA-----GAAGACAGAGGTTTGTCACTGTGG- 3799
 Db 3693 GATGGAAGAACCAAGAGATGACATTTTTCGAAGACTTGAAGGTTCTATGACAGAGC 3752
 QY 3800 -----ACCCAAGTCAAAAGCTGCTGATGTGACATATGACAGTGAAGTATGGAGTG 3854
 Db 3753 ATACAGTCACTCGGAAAGAGTCACTGAATTTGCCCTTAAGGTGAAGGACAGAAAGAA 3812
 QY 3855 GCCGGGTGTCAAGAAAGAGAGATCTGAAGTGCAGAGTCTTAG-----C 3899
 Db 3813 GGTGAATGTAAGAAAGATGATGCTCTTGAATGCAAGTCAAGCTTAAGTCTCTCATCC 3872
 QY 3900 CTGAGAGGAGAGATGAGAACTGACGTTGAAGAAAGAAAGAGGACAAAGCAGAG 3959
 Db 3873 CCGGTGAAGAGAGATGTGTCTCAAGTCAAGAGGAGAAACAGAGCAGAGCAACC 3932
 QY 3960 CAAGTGAAGTGAAG---AAGGTGACAGAGAAACAGCCGCTCTGAGCATGAAGAACTAC 4016
 Db 3933 CATGTGAATGAAGAGAGCTTGAAGCAAGAAACAGCTGTACCTGATCTGAAGAGTCA 3992
 QY 4017 GGAAGCCAGTCTTGAACATTTGAATGCACTGACAGAGGAGGAGGAGCACTGGGAAGC 4076
 Db 3993 AAGCAGCTCTTCCAGACAGTGAATGTGCCATCATATGATGAGCAAGAGAGTCAAGCAGT 4052
 QY 4077 CTGGAAGAAACCTTCTC---TCCAGACCAAGACAAAGAGGTGATAGAGTTCAA 4133
 Db 4053 TTGGAAGGAAGCCCTCTCTCTGCTGCTAGGTCAAGAGGAGGAGATGACCAAAATTTAA 4112
 QY 4134 GTTCAAGCTGAGACACAAACAGTCACTCAACAGCAAGAGTGTGAGAAA-----GCTC 4187
 Db 4113 GTTCAAGCTCTGAGGCACTCATTTCACTTAACAGGCGCTGACAGAGAGGAAAGTCTTA 4172
 QY 4188 ATGAAGACGTTGTGATTTCAAGACAGGTGAAGTCCAGAGTGTGATGTCACATTTA 4247
 Db 4173 GGAGAACTGCCCAATTTTGAAGAACAGGTGAAGCTTGAAGGCTGACAGTGCACATTTA 4232
 QY 4248 TTACCAAGTGAAGATCTCTGCAACGSGTGGCACTGGAATCTTACACATGCAAGAGAC 4307
 Db 4233 GTTCTGAAGAGAAATCTCTGAAAATAAGACTTTGCGCTCATCCAGGGGAAGAT 4292
 QY 4308 ACGGTACCCCTGGGGCTGAGTCTCAGCAGAAATCATCCCAATCATATGTAATCTCTCT 4367
 Db 4293 GGTGTGCCACAGGGCCCGACTGTCTAGGCAAAATGACACAGATGATAGTATCTGTACT 4352
 QY 4368 CCGTGAAGCAACCTTACATCTGACCTTACAGAGAGAAATPAAGCATCTCCAGAGAGCGA 4427
 Db 4353 ACCAAGAAAGGCTTAAAGTTTCCGACTTGAAGAGAGAAACCAATCATCTGAAGTGAAG 4412
 QY 4428 TCAAGAGAAAGAGACAAACAGCAGATCTGTCTGATGCTGACGGAAGAGAGTACAGCA 4487
 Db 4413 TCAGATGAAGTGAAGACAGAGTGTGCTTGGC-----AGAGGTCTAAAGTGAAGTATACA 4466
 QY 4488 ATGAAAAAGTCTCAAGGCTGAACCTGAGATCTGGAATCTGGAATGAGATGAAGCAAG 4547
 Db 4467 ATTTG---AGATTTTGAAGCTGGAATGGAATTTTGAAGCTTGAAGCAAGAGCTTAA 4523
 QY 4548 ATTGTGTAAGTCAATTCAGACAGCCGTTTGAACAGTTTGCACGTAAC---AGAAAACAGC 4604
 Db 4524 CTTGTCAAAATCATCATCAAGACAGCCGTTGACAGGTTTGAATGATCAAGAAAGAACAGCC 4583

QY 4605 CCGAAACTCATGCTTATGATTCACAGACCAGGTTCTCTGTCATG 4648
|||||
Db 4584 ACCGAATGTTGAGCTGCTGAGTTACAGACACAAGCTCAGTGATAAAAGCTGACAGCCAG 4643
QY 4649 CAGGCTTTGACAGCAGGAGGCCCAACAGATGCTGGACAAAATGAAAGATGCCAAGATGAA 4708
|||||
Db 4644 GACGCTGACAGGAAACGAGAAAGAGGAGAGAACCTCTGGCCTCTGCACAGATGAA 4703
QY 4709 ACACCCAGTCCGCGAGGCCA-----GAGAGACTTGCAGTCTCT 4747
|||||
Db 4704 ACACCAATTAATCTTCAGCCAAAGAGGAGTCAAGTCAACCGCAGTGGGACACACATCTCT 4763
QY 4748 GACGCTTCTGAGGSCATGGCTCAGCTCGGAATGCTTGCAGGCTTGCAGTGAAGC 4807
|||||
Db 4764 GATATTTCCAAAGACATGAGTGAAGCCTCAGAAAAGACCATGACTGTTGAGGTAGAAGGT 4823
QY 4808 GCGGTGTCAA----- 4818
|||||
Db 4824 TCCACTGTAATGATCAGCAGCTGGAAGAGTCTGCTCCCATCTCGAGGAAGGAGGT 4883
QY 4819 -----AGTAAGCAATT 4928
Db 4884 GGAGCTGGAACAAAGTCTGTCCAGAAAGATGATGCTCATGCTTGTAGCAGAAAGAATA 4943
QY 4829 GAGAGCTGCTCTCAACCCNAGA--TCCAAAGAGGATGCTGCTGATGGCCCTCAG 4986
|||||
Db 4944 GAGAAGTCACTAGTTGAACCGAAGAAAGATGAAGAAAGTGTATGATGACCCCTGAA 5003
QY 4887 CTCAAAGCTTTAGCCAGGAGGAGCCAGTCCCTCTCGAAACCTTAACCAAGAAATCCCCA 4946
|||||
Db 5004 AACGAGAACTCAGCCCTGGCTGATCTGATGCTCAGAGGCTTAACCAAGAGTCCCCA 5063
QY 4947 GACACCAACGGAACAAAGCTAAACGAGAGGGGATCCCCCAAAAGTTGAGGTTCCAGGAA 5006
Db 5064 GATACAAATGGACCAACAAAGAGAGAGGAGATGCCAGGAGTGAAGATTGCAGGAA 5123
QY 5007 GAAGAAATGTTACCAAGTCACTAAGAGNAACAGGCCAGGAGAGAGGAGCTCCAG 5066
Db 5124 GGAAGTGCACATGTAATCAGATAAAGCGATCAACCCCAAGCAGAGGAGGATTAACAG 5183
QY 5067 GAGCAAGGAGGAGCTGGCAGAGATCTTAAGATGTTAGTT----- 5106
Db 5184 AACAGAGAGAGNATCTGCAAGTCAAGACTTACAGAACTTTAAACATCATGCAAGTTA 5243
QY 5107 -GCTCATTTGATCTGTAAAGACAGAAATGTGAACCAAGTCAACAGAACAA--GATGC 5161
Db 5244 AACTCATTTGCTGTTGGAAAGACCAAGATGTGAAGACAGTGTAGTGAAGAAATGAATGC 5303
QY 5162 TGCTGTTGGACCTTGAGACCAAGATTTTCAGAGCCCATGAGATCCAGAGAGAGGCGCGT 5221
Db 5304 TGCTGTGAGA-CTGAAGACCCAGTATTTTCAGA-ACTTTGAGAAATTTGAGAGAGGAGCAGAT 5361
QY 5222 CCAATGATTTCCACCCAGTAGAGACCCGACAACTCTGAGGCTTCACTCGGAGCTAGAG 5281
Db 5362 CACTGATCTCAATTTCTAGAGAGC-CCCTGACAACTCTGAGGCTTCACTAGGAGCTAGAG 5420
QY 5282 CCAGCTAACATTTCTCGTTTCAAGACTGCTGCTTTGATTTGCCCCCTTGATGCGCTCCGTGT 5341
Db 5421 CCATTTAACTTTCTTTCCAGACCAACCTACAAATTTTCCCTTGATTAAC----- 5472
QY 5342 ATTTCTAACATTTCTCGTTTCAAGACTGCTGCTTTGATTTGCCCCCTTGATGCGCTCCGTGT 5401
Db 5473 -----CAPATA 5478
QY 5402 ATTTGCGATTTAAGTCTCGTTTCTCAACTGGAACCA-ATTCTGCATACCTAGTTCC 5460
Db 5479 AATTCGATTTAAGTCTCAAAATCTTAACCTGGAATGAGGTTGGAATACCTAGTTCT 5538
QY 5461 ACTTCTCAAACTGGAGCATCTCTCTTTATGTAATTTATATGATTTTATGATGCTCTCC 5520
Db 5539 GCTTCTGAACTGAGTATCACTTTTACATATTTATATGATTTTATGATTTTATGATG--TCC 5595
QY 5521 TCCTGTACCTATTGATATATTTTTTTCTTAACGTTT-----AAGCACATG 5563

Db 5596 TCCTGATCTATGATATATTTTTTCTTAATGTTTAAGGAAATGTGCAGGATCTCATG 5655
QY 5564 CTTTTTGTATTTATGCAATATATAACGGGTGTGCAGCGCATAGCGACGCTTTTGAAGTCC 5623
Db 5656 CTTTTTGTATCACACAGTATATGATGGGCATGTGCATAGTCGAGGCTTTGGGAGCTTT 5715
QY 5624 AAGCCTCAACTGTAACTGCTGAGCAAAACAGATAACATTCT-----CTGGCAAGAGAG 5674
Db 5716 AAGCCTCAGTTATATAACCCACGAAACAGAGCCTCTAGATGTAAACATTCTCGATCAA 5775
QY 5675 ACAAGCTTTTTTAAAGTTTACTGATCTTAGATCTGTGGGCTTCTAGTCTCTCGAAAGT 5734
Db 5776 GGTACAAATCTTTTAAATTCACATGATTTAGGTCATATTTAGTGGTACTCTCGAAAT 5835
QY 5735 GGTGTTTTCCTATGCACAGCGAGCTCAGAAATAAAACCCCATTTTGAACATCCAGGA 5794
Db 5836 GGTCACTTTCCTATTACACGGAGTGTCTAAACATAAAGCATTTTGAACATACAGAA 5895
QY 5795 TGTCCCAATATTAACATGATTTTTCCTCCCTTTTGTCTTAATCCAGTCCAGGTTGGAAG 5854
Db 5896 TGTCTTATTTGTCATTTGGGAAATTT-----TTCTTTCTAACCCAGTGGAGTTAGAAAG 5948
QY 5855 AAGTCTCTCTGCTGTCAGA-----TTAAGCCCTCTCTTTAATGATAT 5897
Db 5949 AAGTTATATTTCTGGTAGCAAAATTAACCTTTTACATCTTTTCTTATGTTGTTT 6008
QY 5898 GGACAAATGAGTGTGCTTAAGGCCATGAGA-----TGTTTCTCTAATGCA 5941
Db 6009 GGACCAATGAGTGTGCTTAATCTGAGGCAAGTAGTAATGTTTATATGTTATGAA 6068
QY 5942 GAAGGAATCTGTTGATGCTTTTTCCTTCTTCTTCTATGCTGACCGAATTCATATG 6001
Db 6069 GAAAGAAATTTGTTAAG--TTTTTGATTTCTACTTATATGCTGACTGCAATTCACA 6126
QY 6002 CAGATCGAGTGTGCTGCTTCTTTACAGATGTTATTTTATGATAGTACTGAGTTGTCT 6061
Db 6127 TGGCATGAAATTAAGTCAAGTCTTTTACAAATGTTATTTTATGATAGTACTGAGTTGTTT 6186
QY 6062 GTCTTATATCTGTGCCCTTCTTTAAGAACAAATGTTGCAATATGTTTCTTTTGGATAAAT 6121
Db 6187 GTGCCATATTTGTGCCATCTTTTAAAGAACAAATGTTGCAACATTCATTTTGGATAAGTT 6246
QY 6122 GTGATTTGACAACTGATTTTAAATAAA 6147
Db 6247 GTGATTTGACGACTGATTTTAAATAAA 6272

RESULT 15

AL590413

LOCUS

DEFINITION

Human DNA sequence from clone RP11-351K16 on chromosome 6, complete

sequence.

ACCESSION

AL590413

VERSION

AL590413.18 Gi:17426497

KEYWORDS

HTG.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 104939)

AUTHORS

Martin,S.

TITLE

Direct Submission

JOURNAL

Submitted (06-DEC-2001) Wellcome Trust Sanger Institute, Hinxton,

Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

humayy@sanger.ac.uk

On Dec 8, 2001 this sequence version replaced gi:17384482.

During sequence assembly data is compared from overlapping clones.

Where differences are found these are annotated as variations

together with a note of the overlapping clone name. Note that the

variation annotation may not be found in the sequence submission

corresponding to the overlapping clone, as we submit sequences with

only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL, SW, SWISSPROT, Tr, TREMBL, Wp, WORMPEP. Information on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contig of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at

<http://www.sanger.ac.uk/HGP/Chr6>
<http://www.sanger.ac.uk/HGP/Chr6> RPI1-351K16 is from the library RPI1-11.2 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: PBAC3.6
 IMPORTANT: This sequence is not the entire insert of clone RPI1-351K16. It may be shorter because we sequence overlapping sections only once, except for a short overlap.
 The true right end of clone RPI1-351K16 is at 104939 in this sequence. The true left end of clone RPI1-108N8 is at 59858 in this sequence. The true right end of clone RPI1-403M6 is at 2000 in this sequence.

FEATURES

source

1. 104939
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="6"
 /clone="RPI1-351K16"
 /clone_1b="RPI1-11.2"
 51026..51045
 /note="Sequence from AC023201 sequenced by WIBR."

misc_feature
 51799..51952
 /note="Single clone region. Sequence from reads from a short insert library derived from a single pUC clone. Restriction digest data confirm the assembly."

misc_feature
 52345..52447
 /note="Sequence from reads from a short insert library derived from a single pUC clone. Restriction digest data confirm the assembly."

misc_feature
 86009..86146
 /note="Sequence from overlapping clone RPI1-108N8 (AL590543). Assembly confirmed by restriction digest."

misc_feature
 87770..87807
 /note="Sequence from uni-directional dGTP big dye terminator reads only."

BASE COUNT 30065 a 22273 c 22882 g 29719 t

ORIGIN

Query Match 30 9%; Score 1906.4; DB 9; Length 104939;
 Best Local Similarity 68.2%; Pred. No. 0; Mismatches 1246; Indels 186; Gaps 23;
 Matches 3070; Conservative 0;

297 GTTGACAGCAGAGTCAGAGATGTAGAGAAAAGACCGAGTTGAAGAAATGGCGCC 356
 Db 9415 GTTGGACAGAGACTCTGAGATGTAGAGAAAAGACCTCGATTAAGAGATGGCTACT 9474
 Qy 357 AACTCCAGAGCTGTGAAGATATCAAAAGATGGCGAGAGACATCAAAATTAATT 416
 Db 9475 AAGTCACGGTTGTTCACGACATCAAGATGAGGCGAGAGAGACACCCGAAATTAATC 9534
 Qy 417 GAACAGATCCCTGCTTCAGAAAACATGTGAGAGAAATGTTAGCTGTGAGTCCAG 476
 Db 9535 GAACAGATTCCTTCTTCAGAAACAAATTAAGAGAGCTTAACACACCTGAGTCCAG 9594
 Qy 477 GCTAATGATGTTGGCTTCAAGAAATTAATTTGTTGTTTAAATTCACGCTGAAG 536
 Db 9595 GCTAATGATGTTGGATTTAAGAAAGCTTTAAGTTCTTGGCTTAAATTCACGCTGAAG 9654

Qy 537 AAGATAAATAATGAAAAATGCAGATCTGTCCAACTACTACTGTCAAGAGATGAAGCC 596
 Db 9655 AAGATAAAGACAGAGAAAGCCGACACTGTCCAGCTACTCTGTGAAGAAAGATGAAGGG 9714
 Qy 597 GAAGGGCAGAGCCTCTGTGAGCTGAGAGACCAAGAGACCCACTGTGAGACTGCC 656
 Db 9715 GAGGAGCAGCA-----GGGGCTGGGCAACCAAGAGACCCACTGTGGGGCT--- 9762
 Qy 657 GTGGAGGTGAGATCCAAAGAAAGTGAAGTGAAGCAATCCACAGAGAAAGCAAGAGCC 716
 Db 9763 ---GAGAGAGCAGATCCAAAGAAAGCAAGCAACCAAGCAATTCAGAGAAACCGAGAG 9819
 Qy 717 ACCCTGAAGCAAGACAGAGCAGACAGAAATCCCTTCAGAGCCGATGTATCAAGCG 776
 Db 9820 ACCCTGAAGCTGAGCAAGACAGAGCAAGAAATTTCTCCAGCGCATGTGGCA--- 9876
 Qy 777 GCTGAGAGAGGCCAAGATGAAAGAGAGAAACAAAGAGAGCCCAAGTCC 836
 Db 9877 GCAGTGAGAGAAATCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGTCT 9936
 Qy 837 CCAGAAATCCCGAGACGCCGCTGCAAGTGAACATCTCTTCAAGAGTCTTC 896
 Db 9937 GCAGAAATCTCGACTAGTCCGTGACAGTGAACAGATCAACCTTCAAAAATTTCTTC 9996
 Qy 897 ACTCAGGTTGGCGCGCTGCGCAGAGACAGCTTCAAGAAATCAAAAGAGATGAT 956
 Db 9997 ACTCAGGTTGGCGCGCTGCGCAGAGACAGCTTCAAGAAATCAAAAGAGATGAT 10056
 Qy 957 CTGGAATCTGCCAG 1016
 Db 10057 GTGGAATCTTGAAG 10116
 Qy 1017 GAAAAG 1049
 Db 10117 GAAAAG 10176
 Qy 1050 CCGGAGAGAGACACAG 1109
 Db 10177 CCGGAG 10236
 Qy 1110 TTGGAAGACAGAGTTGTGACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1169
 Db 10237 TCAG 10296
 Qy 1170 ACGGAAGGTTTGAAG 1226
 Db 10297 ACGGAAGGTTTGAAG 10356
 Qy 1227 GTGAGCAGCTGAG 1286
 Db 10357 GTGAGCAGCTGAG 10392
 Qy 1287 GTGAGGTTGAAG 1346
 Db 10393 ACGGAGGTTGAAG 10452
 Qy 1347 GTCCCCCAG 1406
 Db 10453 GAACCTCAG 10512
 Qy 1407 GGAGAGAGACACACTCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1466
 Db 10513 GGAG 10572
 Qy 1467 CCAG 1526
 Db 10573 CCAG 10632
 Qy 1527 GGAAGTCTTGAAG 1586
 Db 10633 GGAAGTCTTGAAG 10692
 Qy 1587 CAGAAGGGAG 1646

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 12, 2002, 14:13:30 ; Search time 1074 Seconds
(without alignments)
12916.489 Million cell updates/sec

Title: US-09-902-432-3
Perfect score: 6160
Sequence: 1 CGGGCGGGGGGAGTGAAG.....aataaacatttgaactac 6160

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: N_Geneseq_101002.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
25: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5182.4	84.1	5200	18 AAV02302	Rat tumour suppress
2	4696.6	76.2	5074	18 AAV02301	Rat tumour suppress
3	2179.6	35.4	6614	24 ABNS59861	Novel human coding
4	2144.4	34.8	6608	24 ABNS96942	Gene #3440 used to
5	2144.4	34.8	6608	24 ABL61938	Colon adenocarcino
6	2118.8	34.4	6605	19 AAV23545	Human gravin polyp
7	2118.8	34.4	6605	21 AAV74903	Human gravin codin
8	1920.6	31.2	6886	23 AAS95205	DNA encoding novel
9	550	8.9	563	24 ABK63316	Rat sequence diffe

10	172.2	2.8	377	20	AA040328	Human secreted pro
11	162.2	2.6	178	24	ABK63568	Rat sequence diffe
12	105.2	1.7	483	20	AAV86620	EST clone BF146.
13	101.8	1.7	459	23	ABV56830	Human prostate exp
14	95.4	1.5	400	22	AAI86887	Human polynucleoti
15	92	1.5	339	20	AAV86330	EST clone AJ169.
16	81.8	1.3	10732	21	AAAI0594	Gene encoding a su
17	79.4	1.3	385	22	AAI83958	Human polynucleoti
18	78.6	1.3	3399	17	AAAT05868	Chicken leucocytoz
19	76	1.2	49999	20	AAZ23891	Murine LOBO genom
20	76	1.2	49999	20	AAZ23896	Murine LOBO homolo
21	72.8	1.2	2800	22	AAZ07131	Canine retinitis p
22	72.4	1.2	1886	16	AAQ87587	DNA encoding 'leuco
23	72.2	1.2	6741	21	AAAI0595	Gene encoding a su
24	71	1.2	575	22	ABA50472	Human breast cell
25	71	1.2	575	22	ABA68422	Human foetal liver
26	71	1.2	575	22	ABA35413	Probe #13879 for g
27	71	1.2	575	22	AAK16793	Human brain expres
28	71	1.2	575	22	AAK42567	Human bone marrow
29	71	1.2	575	22	AAI23315	Probe #13248 for g
30	71	1.2	575	22	AAI48636	Probe #17322 used
31	71	1.2	575	22	AAI08956	Probe #8947 used t
32	71	1.2	575	24	ABS16615	Human genome-deriv
33	71	1.2	1969	22	ABA45341	Human breast cell
34	71	1.2	1969	22	ABA55830	Human foetal liver
35	71	1.2	1969	22	ABA25506	Probe #3972 for ge
36	71	1.2	1969	22	AAK04048	Human brain expres
37	71	1.2	1969	22	AAK29533	Human bone marrow
38	71	1.2	1969	22	AAI14105	Probe #4038 for ge
39	71	1.2	1969	22	AAI35486	Probe #4172 used t
40	71	1.2	1969	22	AAI03958	Probe #3949 used t
41	71	1.2	1969	24	ABS04084	Human genome-deriv
42	70.8	1.1	799	19	AAV55831	Nucleotide sequenc
43	70.8	1.1	1926	21	AAAS0254	Epstein Barr virus
44	70.8	1.1	1926	22	AAK82902	EBV tethering prot
45	70.8	1.1	2580	21	AAK75454	Nucleotide sequenc

ALIGNMENTS

RESULT 1	
AAV02302	ID AAV02302 standard; cDNA; 5200 BP.
XX	AC AAV02302;
XX	AC AAV02302;
DT	20-JUL-1998 (first entry)
XX	Rat tumour suppressor gene SSeCKs.
DE	XX SSeCKs; tumour suppressor gene; rat; protein kinase C; mitosis;
KW	cancer; malignancy; cell proliferation; Alzheimer's disease;
KW	therapy; ss.
OS	Rattus sp.
XX	XX Key Location/Qualifiers
FT	CDS 1..4791
FT	/*tag= a
XX	XX
PN	WO9740059-A1.
XX	XX
PD	30-OCT-1997.
XX	XX
PF	18-APR-1997; 97WO-US06830.
XX	XX
PR	18-JUN-1996; 96US-0665401.
PR	19-APR-1996; 96US-0635121.
XX	XX (GELM/) GELMAN I.
PA	(JAKE/) JAKEN S.
XX	XX

PI Gelman I, Jaken S;
XX
XX WPI, 1997-535770/49.
DR P-PSDB; AAMJ347.
XX
XX
PT Tumour suppressor gene SSeCKs - used as a mitotic regulator, and
inhibitor of malignant phenotype
XX
XX
PS Claim 1, Fig 11A-L; 162bp; English.
XX
XX This rat cDNA sequence codes for the full-length form of the novel
CC tumour suppressor protein SSeCKs (see AAMJ347). The SSeCKs gene
CC product is a substrate of protein kinase C and acts as a negative
CC regulator of mitosis and as an inhibitor of the transformed
CC phenotype. A partial SSeCKs clone (see AAO0230) was identified by
CC searching for cDNAs whose abundance was low in NIH 3T3 cells and
CC decreased following the expression of the activated oncogene v-src.
CC A 5'RACE product was spliced to this truncated sequence to construct
CC the full-length SSeCKs cDNA sequence. This full-length SSeCKs
CC nucleic acid sequence, as well as homologous and hybridising nucleic
CC acids are claimed, as are vectors comprising such nucleic acids,
CC encoded proteins, host cells and methods of inhibiting the
CC expression of a transformed phenotype in a host cell by introducing
CC the nucleic acid. Introduction of a SSeCKs nucleic acid or gene
CC product into a host cell inhibits mitosis of the host cell,
CC allowing the treatment of diseases associated with disorders of
CC proliferation and/or with the expression of a malignant phenotype.
CC SSeCKs can also be used to treat or identify disorders of
CC cytoskeletal structure and cellular architecture (such as
CC Alzheimer's disease), and may be a marker for aberrancies in
CC fertility and/or nervous system development.

Sequence 5200 BP; 1612 A; 1201 C; 1563 G; 824 T; 0 other;

Query Match	84.1%;	Score 5182.4;	DB 18;	length 5200;
Best Local Similarity	99.8%;	Pred. No. 0;		
Matches 5189; Conservative	0;	Mismatches 11;	Indels 0;	Gaps 0;

OY	33	ATGGCGGAGGCAAGTTTCCACCGAAGCGGGAGACCCCGACGACGCGGGGGAGAGACAG	92
Db	1	ATGGGCGGAGGAGTTTCCACCGAAGCGGGAGACCCCGACGACGCGGGGGAGAGACAG	60
OY	93	CCGAGCGAGCTGGTCTCAGTGGCCATGGGCGCAGTGTGAAGCTTCGGAGCAGCTGGA	152
Db	61	CCGAGCGAGCTGGTCTCAGTGGCCATGGGCGCAGTGTGAAGCTTCGGAGCAGCTGGA	120
OY	153	GACCCCCGGAGCGGGAGACCCCGCACCAAGTCTCCACAGAAAGATGGCCAGCTGTCTT	212
Db	121	GACCCCGCGAGCGGGAGACCCCGCACCAAGTCTCCACAGAAAGATGGCCAGCTGTCTT	180
OY	213	GTCAACGGCGTAGCTGAAACAAGAGATGTCCATGTCCAAAGSAAAAACAAGAGGGGAG	272
Db	181	GTCAACGGCGTAGCTGAAACAAGAGATGTCCATGTCCAAAGSAAAAACAAGAGGGGAG	240
OY	273	GAGGAAGAAGTCCTTGATGAGAGATGTTGGACAAGGAGAGTCACAGAAGTGTGAGAGAAAA	332
Db	241	GAGGAAGAAGTCCTTGATGAGAGATGTTGGACAAGGAGAGTCACAGAAGTGTGAGAGAAAA	300
OY	333	GACCGAGTTGAAGAATAAGCGGCGCAACTCCACAGCTGTGAAGATATCACAAAGAGATGG	392
Db	301	GACCGAGTTGAAGAATAAGCGGCGCAACTCCACAGCTGTGAAGATATCACAAAGAGATGG	360
OY	393	CAGGAGGAGACATCAGAAATAATGAAACAGATCCCTGCTTCAAGAAACAATGTGGAAAGAA	452
Db	361	CAGGAGGAGACATCAGAAATAATGAAACAGATCCCTGCTTCAAGAAACAATGTGGAAAGAA	420
OY	453	ATGTGACAGCCGTGAGATGCCAGGCTTAATATGTGGCTTCAAGAAAGATTTAAATTT	512
Db	421	ATGTGACAGCCGTGAGATGCCAGGCTTAATATGTGGCTTCAAGAAAGATTTAAATTT	480
OY	513	GTGTGTTTTAAATTCACGGTGAAGAAGATAAAAATGAAAGTCAGATCTGTCCACTA	572
Db	481	GTGTGTTTTAAATTCACGGTGAAGAAGATAAAAATGAAAGTCAGATCTGTCCACTA	540

QY	573	CTCTCTGTCAAGAAAGATGAAGCCGAAAGGGCGAAGAAAGCTCTGTCTGGAGCTGGAGACCAAC	632
Db	541	CTCACTGTCAAGAAAGATGAAGGGCGAAGGGGCGAAGAGCCTCTGTCTGGAGCTGGAGACCAAC	600
QY	633	CAGAGCCCAAGTGTGGAGACTGCGCTCGAGAGAGTCAGATCCAAAGAAAGTAGACTGAAG	692
Db	601	CAGAGCCCAAGTGTGGAGACTGCGCTCGAGAGAGTCAGATCCAAAGAAAGTAGACTGAAG	660
QY	693	CAATTCACAGAGAAACAGAAAGGCACCTTGAAGCAAGAAACAGAGCAGCACAGAAATCCCC	752
Db	661	CAATTCACAGAGAAACAGAAAGGCACCTTGAAGCAAGAAACAGAGCAGCACAGAAATCCCC	720
QY	753	CTTCAAGCCCAATTTGATCAAGGGCTGGAGAAAGACCCAAAGATGAAGGAAGAAAGAAA	812
Db	721	CTTCAAGCCCAATTTGATCAAGGGCTGGAGAAAGACCCAAAGATGAAGGAAGAAAGAAA	780
QY	813	CAAGAGAAAGAGCCCAACMAATCCCGAGAAATCCCGAGCAGCCCAAGTCAACAGTGAGACA	872
Db	781	CAAGAGAAAGAGCCCAACMAATCCCGAGAAATCCCGAGCAGCCCAAGTCAACAGTGAGACA	840
QY	873	AACATCTTCCTTCAAGAAATTCCTCACTCAAGGTGGGCGGCTGGCGCAAGAACCAAGC	932
Db	841	AACATCTTCCTTCAAGAAATTCCTCACTCAAGGTGGGCGGCTGGCGCAAGAACCAAGC	900
QY	933	TTCAAGAAATCAAAAGAGATGATCTTGAAACTGCCGAGAGAGAAAGAGCAAGAGCA	992
Db	901	TTCAAGAAATCAAAAGAGATGATCTTGAAACTGCCGAGAGAGAAAGAGCAAGAGCA	960
QY	993	GAAAAATGATCCAGGAAAGAAAGAAAGAAAGACAGAGCCAGCTCCGAGAGACACAGAGCG	1052
Db	961	GAAAAATGATCCAGGAAAGAAAGAAAGAAAGACAGAGCCAGCTCCGAGAGACACAGAGCG	1020
QY	1053	GCAGAGACACAGACCAGGCCAGAGTTGTGACAGACACTACAGAAAGTGGAGCTGCTTTG	1112
Db	1021	GCAGAGACACAGACCAGGCCAGAGTTGTGACAGACACTACAGAAAGTGGAGCTGCTTTG	1080
QY	1113	GAAAGCCAGGTTGGTGACTCTGAGGGCATGTCAGAGAGAGAGTGTCTCTTTGGCAAG	1172
Db	1081	GAAAGCCAGGTTGGTGACTCTGAGGGCATGTCAGAGAGAGAGTGTCTCTTTGGCAAG	1140
QY	1173	GAAGTGTTGATGAGAAAGATGGAAGGCCCAACAAAGAGTTGTTGCAAGAGTCCACGTGAGC	1232
Db	1141	GAAGTGTTGATGAGAAAGATGGAAGGCCCAACAAAGAGTTGTTGCAAGAGTCCACGTGAGC	1200
QY	1233	ACCGTGGAGAAACAGAGAGAGAGACAGAGAGAGAGAGAGAGCTGAAGGGGCGTGGTG	1292
Db	1201	ACCGTGGAGAAACAGAGAGAGAGACAGAGAGAGAGAGAGAGCTGAAGGGGCGTGGTG	1260
QY	1293	GTAGAAGGAACAGGAGAAATCTTTGCCCTCTGAGAAACTGGCTGAGCCCCAGAGAGTCCCC	1352
Db	1261	GTAGAAGGAACAGGAGAAATCTTTGCCCTCTGAGAAACTGGCTGAGCCCCAGAGAGTCCCC	1320
QY	1353	CAGGAAGCTGAGCTGCTGTGAGGAGCTGATGAAGAGCAGAGAGATGTGTCTCTGGAGGA	1412
Db	1321	CAGGAAGCTGAGCTGCTGTGAGGAGCTGATGAAGAGAGATGTGTCTCTGGAGGA	1380
QY	1413	GACCACTCAACTGACAGAGCTTAAGTCCCTGAAGAGAAAGCTGCCCAAAACACCCAGAA	1472
Db	1381	GACCACTCAACTGACAGAGCTTAAGTCCCTGAAGAGAAAGCTGCCCAAAACACCCAGAA	1440
QY	1473	GGCATTGTCAAGTGAAGTGAAGATGCTGCTCTCAGAGAAAGAAATCAAGTACAGGGAAGT	1532
Db	1441	GGCATTGTCAAGTGAAGTGAAGATGCTGCTCTCAGAGAAAGAAATCAAGTACAGGGAAGT	1500
QY	1533	CCCTTGAAGAAACTCTCAAGTACAGGCTTTAAGAAAGCTGTCTGGAAAGAACAGAG	1592
Db	1501	CCCTTGAAGAAACTCTCAAGTACAGGCTTTAAGAAAGCTGTCTGGAAAGAACAGAG	1560
QY	1593	GGGAAACGAGAGAGTGGGGGAGACGAAAGAGCTTGGAGAAATACCAACATTTCAACCCGA	1652
Db	1561	GGGAAACGAGAGAGTGGGGGAGACGAAAGAGCTTGGAGAAATACCAACATTTCAACCCGA	1620

QY 1653 TCCCAGAGAGTCTGATGACAGAGGAGAGAGTCTGCGTCTGTCCTCCCGAGAGCGCT 1712
Db 1621 TCCCAGAGAGTCTGATGAGCAGAGGAGAGAGTCTGCGTCTGTCCTCCCGAGAGCGCT 1680
QY 1713 GAGGAGACCACTCTCTGGAGAAAGGCGCTGGAAGCACCCACAGGATGGGAAAGCTGAG 1772
Db 1681 GAGGAGACCACTCTCTGGAGAAAGGCGCTGGAAGCACCCACAGGATGGGAAAGCTGAG 1740
QY 1773 GAAGGAATCTACTCCGATGGAGAGAAGAAGAGAGAGGATCACTCTCCCTGGGCACTCTTC 1832
Db 1741 GAAGGAATCTACTCCGATGGAGAGAAGAAGAGAGAGGATCACTCTCCCTGGGCACTCTTC 1800
QY 1833 AAAAGATGGTGACACCCNAGAAACGGTCCGAGACCTCTTGAGAGTGCACAGAGAGAA 1892
Db 1801 AAAAGATGGTGACACCCNAGAAACGGTCCGAGACCTCTTGAGAGTGCACAGAGAGAA 1860
QY 1893 GAGCTGGAGAAAGTCAAGAGGCGCACCTTGCTCTCACTGATAGCACAGTGTCAAGAATG 1952
Db 1861 GAGCTGGAGAAAGTCAAGAGGCGCACCTTGCTCTCACTGATAGCACAGTGTCAAGAATG 1920
QY 1953 CAAGATGAAGTCAAACTGTTGGAGGAAACAAAGCCAGAGGAAACAAAGCGTAGGGTG 2012
Db 1921 CAAGATGAAGTCAAACTGTTGGTAGGAAACAAAGCCAGAGGAAACAAAGCGTAGGGTG 1980
QY 2013 GATACCTTCACTGCTTGGGAAGCACTGATTTGTTGTCGATCATCAAGAGAGAGCAAGG 2072
Db 1981 GATACCTTCACTGCTTGGGAAGCACTGATTTGTTGTCGATCATCAAGAGAGAGCAAGG 2040
QY 2073 AAGCATCTCTTTCAGATGATGAAGGAGGCGCAAGGACACTGGGAGGGGACAGTCAAGA 2132
Db 2041 AAGCATCTCTTTCAGATGATGAAGGAGGCGCAAGGACACTGGGAGGGGACAGTCAAGA 2100
QY 2133 GCAGAGAGGCGCAAGAGCAAAAGAGCGGAAACAGACGCTGTTCTTCGAGACCCAG 2192
Db 2101 GCAGAGAGGCGCAAGAGCAAAAGAGCGGAAACAGACGCTGTTCTTCGAGACCCAG 2160
QY 2193 GAGCAGACCAAGCGCAAGGAGTTCTTCAACCAGCGAGCGGGAAGCCCTTCGAAAGG 2252
Db 2161 GAGCAGACCAAGCGCAAGGAGTTCTTCAACCAGCGAGCGGGAAGCCCTTCGAAAGG 2220
QY 2253 GAAGGTGCTCCACTTGGGAGTCAATTTAAAGATTAGTCACTCCAAGAAAAATCCAAG 2312
Db 2221 GAAGGTGCTCCACTTGGGAGTCAATTTAAAGATTAGTCACTCCAAGAAAAATCCAAG 2280
QY 2313 TCRAAATCGAAGAGAAAGCCGAGACCTCTAGTGTAGAGCAGTTGTCCACTGAGATCGAA 2372
Db 2281 TCRAAATCGAAGAGAAAGCCGAGACCTCTAGTGTAGAGCAGTTGTCCACTGAGATCGAA 2340
QY 2373 CCGAGTAGAGAGAAATCTTGGGTTTCCATTAAAGAAATTCATCCCGGACGCGGAGAA 2432
Db 2341 CCGAGTAGAGAGAAATCTTGGGTTTCCATTAAAGAAATTCATCCCGGACGCGGAGAA 2400
QY 2433 AGGCGAGCGGAGCAGAGAACACCTGTGGAAGCTCAGGCGCAGTGGAGATAAT 2492
Db 2401 AGGCGAGCGGAGCAGAGAACACCTGTGGAAGCTCAGGCGCAGTGGAGATAAT 2460
QY 2493 GAGCAGACCCCTAATGTCCACCGCTGCTGCTCTGTCTGAGTATAATGCAAGTGGAGG 2552
Db 2461 GAGCAGACCCCTAATGTCCACCGCTGCTGCTCTGTCTGAGTATAATGCAAGTGGAGG 2520
QY 2553 GAGAAGATGGAGCCCGAGGAATACGAGGCTGCCAGCTGTGGGGCTGTGTAAGTG 2612
Db 2521 GAGAAGATGGAGCCCGAGGAATACGAGGCTGCCAGCTGTGGGGCTGTGTAAGTG 2580
QY 2613 TCCGAGAGGCTCACTAGACTCTGTCACACTGTGAGTGTGCGAGTCAATGTGGGACC 2672
Db 2581 TCCGAGAGGCTCACTAGACTCTGTCACACTGTGAGTGTGCGAGTCAATGTGGGACC 2640
QY 2673 AGGCGAGTCAACAGTGTGGAAGCGGCTCTCTTCTGTTGATATCCGCTTCCGTAACAGAA 2732
Db 2641 AGGCGAGTCAACAGTGTGGAAGCGGCTCTCTTCTGTTGATATCCGCTTCCGTAACAGAA 2700
QY 2733 CCTTTGAACACACAGCGGGAGAAAGCCATGCCCTGTTGAAGAGGTCACTGAAAAAGAC 2792

Db 2701 CCTTTGAACACACAGCGGAGAAAGCCATGCCACTCTGTTGAAGAGGTCACTGAAAAAGAC 2760
QY 2793 ATCATTTGCAAGAAACTCTCTGTGCTCACCCAGACGTTTACAGAGGGTAAAGATGCCCAT 2852
Db 2761 ATCATTTGCAAGAAACTCTCTGTGCTCACCCAGACGTTTACAGAGGGTAAAGATGCCCAT 2820
QY 2853 GACGACATGTCTACCAAGTGAAGTGGATTTCACTTCAGAAAGCTGTGACGACACAGAGACC 2912
Db 2821 GACGACATGTCTACCAAGTGAAGTGGATTTCACTTCAGAAAGCTGTGACGACACAGAGACC 2880
QY 2913 TCAGAGGCTCTCGTACTGTAAGAGTTACCGAAGCATCGGGGCCCGAAGAGACACAGAC 2972
Db 2881 TCAGAGGCTCTCGTACTGTAAGAGTTACCGAAGCATCGGGGCCCGAAGAGACACAGAC 2940
QY 2973 ATGTTGTCGACAGTTTCCAGCTGACTCTCCAGACACACAGAGGAAGCCACCCCA 3032
Db 2941 ATGTTGTCGACAGTTTCCAGCTGACTCTCCAGACACACAGAGGAAGCCACCCCA 3000
QY 3033 GTTACGAGAGTAGAGTGTGTCTAGATACAGAAAGAGGAGCGCCAGACGACGCGC 3092
Db 3001 GTTACGAGAGTAGAGTGTGTCTAGATACAGAAAGAGGAGCGCCAGACGACGCGC 3060
QY 3093 ATCTTCCAAGCCGTTGCAGACAAAGTGAAGAGAGTCCAGGTGCTCTGCAACCCAGACT 3152
Db 3061 ATCTTCCAAGCCGTTGCAGACAAAGTGAAGAGAGTCCAGGTGCTCTGCAACCCAGACT 3120
QY 3153 GTGCAGAGAAACGGGTCGAGAGCACTGGAGAAAGTTCGAGAGGTAGAGGAGGACTCCGAA 3212
Db 3121 GTGCAGAGAAACGGGTCGAGAGCACTGGAGAAAGTTCGAGAGGTAGAGGAGGACTCCGAA 3180
QY 3213 GTGCTGCTTCCGAGAAAGAGAGGAGCTTATGCCGAAAGGACCCGTCGAGGAGCTGGA 3272
Db 3181 GTGCTGCTTCCGAGAAAGAGAGGAGCTTATGCCGAAAGGACCCGTCGAGGAGCTGGA 3240
QY 3273 GCTGAGCATCTTGCACAGGCTCTGAGACTGGACAGCTTACTCCAGAGAGCTTGAAGTT 3332
Db 3241 GCTGAGCATCTTGCACAGGCTCTGAGACTGGACAGCTTACTCCAGAGAGCTTGAAGTT 3300
QY 3333 CCTGAAGTCAAGCAGATGTAGACCATGTGCCAGCTGCGAGGTTATCAAGCTCCAGCAG 3392
Db 3301 CCTGAAGTCAAGCAGATGTAGACCATGTGCCAGCTGCGAGGTTATCAAGCTCCAGCAG 3360
QY 3393 CTGATGGAACAGGCGCTGGCCCTCTGAGTCTATCCGAAACCTTGCAGACAGTGAACAAAT 3452
Db 3361 CTGATGGAACAGGCGCTGGCCCTCTGAGTCTATCCGAAACCTTGCAGACAGTGAACAAAT 3420
QY 3453 GGAAGCACTCCCTTAGCAGATTGACACATCTGCAGATGGGACACAGCAAGATGAACCAAT 3512
Db 3421 GGAAGCACTCCCTTAGCAGATTGACACATCTGCAGATGGGACACAGCAAGATGAACCAAT 3480
QY 3513 GACAGCCAGACAGTAAGCCACTGCAGCTGTGAGGAGTCAAGGAGTCAAGAGAGAGAG 3572
Db 3481 GACAGCCAGACAGTAAGCCACTGCAGCTGTGAGGAGTCAAGGAGTCAAGAGAGAGAG 3540
QY 3573 CGGCTACTCTCTCAGAAAGAGGAGCTTCCAGACTACTCTTAATAATGTTCCAGGCCAGGAA 3632
Db 3541 CGGCTACTCTCTCAGAAAGAGGAGCTTCCAGACTACTCTTAATAATGTTCCAGGCCAGGAA 3600
QY 3633 GAACATGGGGAAGAACACAGGAAGAGATGTTCTTGAACCTTACAGCAAGAGTCTTGCT 3692
Db 3601 GAACATGGGGAAGAACACAGGAAGAGATGTTCTTGAACCTTACAGCAAGAGTCTTGCT 3660
QY 3693 GCAGCCGTCGCTTCTGCAAGAGACTGAGGTGCTCAAGGAGTGAAGTGTGCTGCTG 3752
Db 3661 GCAGCCGTCGCTTCTGCAAGAGACTGAGGTGCTCAAGGAGTGAAGTGTGCTGCTG 3720
QY 3753 GATGGGAAAAAGTCAAGAAAGAACAGGAGTGTGTTGTAACCTCTGACCCCAACAGTCAA 3812
Db 3721 GATGGGAAAAAGTCAAGAAAGAACAGGAGTGTGTTGTAACCTCTGACCCCAACAGTCAA 3780
QY 3813 AAGGCTGCTGATGTGACATATGACGTGAAGTGTGGAGTGGCGGCTGTGAGGAAAAAG 3872

```

Db 3781 AAGCTGCTGATGACATATGACAGTGAAGTATGGAGTGGCCGGGTGTCAGAAAAG 3840
Qy 3873 GAGAGTACTGAGTGCAGAGTCTTAAGCTGAGAGGAGAGAGATGAAACTGACCTTGA 3932
Db 3841 GAGAGTACTGAGTGCAGAGTCTTAAGCTGAGAGGAGAGATGAAAAGTGAAGTGA 3900
Qy 3933 AAGGAGAAAAGGAGACAAAGCCAGACAGTGAAGTGAAGAGTGAAGAGAGAGAAACACC 3992
Db 3901 AAGGAGAAAAGGAGACAAAGCCAGACAGTGAAGTGAAGAGTGAAGAGAGAAACACC 3960
Qy 3993 GCTCCCTGAGCTGAAGAAACCTACCGGAAGCCAGTCTGACACTTGATGCGCAGCTCA 4052
Db 3961 GCTCTGAGCTGAAGAAACCTACCGGAAGCCAGTCTGACACTTGATGCGCAGCTCA 4020
Qy 4053 GAGAGGGGAAAGGCACTGGGAAAGCCTTGGAGAGAGCCCTTCTCTCCAGACCAAGCAA 4112
Db 4021 GAGAGGGGAAAGGCACTGGGAAAGCCTTGGAGAGAGCCCTTCTCTCCAGACCAAGCAA 4080
Qy 4113 GCAGGTGCTATGAGTTCATGTTCAAAAGCTTGACACAAAGTCACTCAACACAGAGA 4172
Db 4081 GCAGGTGCTATGAGTTCATGTTCAAAAGCTTGACACAAAGTCACTCAACACAGAGA 4140
Qy 4173 GCTGTGAAAAAGGTGATGAAACGGTGTGATTTGAGAGACAGTGAAGTCCAGAGTGT 4232
Db 4141 GCTGTGAAAAAGGTGATGAAACGGTGTGATTTGAGAGACAGTGAAGTCCAGAGTGT 4200
Qy 4233 GTAGGTGCACTTATTAACAGCTGAGAAAGTCTCTGCAACGGGTGGCCACTGACTCTT 4292
Db 4201 GTAGGTGCACTTATTAACAGCTGAGAAAGTCTCTGCAACGGGTGGCCACTGACTCTT 4260
Qy 4293 CAGCTATGAGAGACACGGTACCCCTGGGGCCTGAGTCTCAGAGCAAGTCCATCCCATTC 4352
Db 4261 CAGCTATGAGAGACACGGTACCCCTGGGGCCTGAGTCTCAGAGCAAGTCCATCCCATTC 4320
Qy 4353 ATAGTAATCTCTGCTCTGTAAGAACCCCTACATCTCTGACCTCAAGAGAAATAAGCGCA 4412
Db 4321 ATAGTAATCTCTGCTCTGTAAGAACCCCTACATCTCTGACCTCAAGAGAAATAAGCGCA 4380
Qy 4413 TCCCGAGAGAGCCGATCAGAGAGAGAGACAAAGCAGATGCTGTCTCTGATGCTGACGGC 4472
Db 4381 TCCCGAGAGAGCCGATCAGAGAGAGAGACAAAGCAGATGCTGTCTCTGATGCTGACGGC 4440
Qy 4473 AAGGAGATGACGCAATGCAAAAAGTCTCAAGGCTGAACCTGAAGTCTCTGAATCTTGAG 4532
Db 4441 AAGGAGATGACGCAATGCAAAAAGTCTCAAGGCTGAACCTGAAGTCTCTGAATCTTGAG 4500
Qy 4533 AGTAGAGCAACAAGTGTGTGTAACGTCATTCAGACAGCGTGTGACAGTTCGACAGT 4592
Db 4501 AGTAGAGCAACAAGTGTGTGTAACGTCATTCAGACAGCGTGTGACAGTTCGACAGT 4560
Qy 4593 ACAGAAACAGCCCGGAAACCTCATGCTTATGATTCACAGACCCAGGTTCTGATGACAG 4652
Db 4561 ACAGAAACAGCCCGGAAACCTCATGCTTATGATTCACAGACCCAGGTTCTGATGACAG 4620
Qy 4653 CTGGACAGCAGGAGCCCAACAGATGCTGACAAAATGAAGATGCCAAGTGAAGAAC 4712
Db 4621 CTGGACAGCAGGAGCCCAACAGATGCTGACAAAATGAAGATGCCAAGTGAAGAAC 4680
Qy 4713 CCAGTCCGACAGCCCAAGAGAGACTTGAAGTCTGACAGTCTGAGAGCATGAGGCTCAG 4772
Db 4681 CCAGTCCGACAGCCCAAGAGAGACTTGAAGTCTGACAGTCTGAGAGCATGAGGCTCAG 4740
Qy 4773 CCTCGAAATGCTTCCGCGCTTGGAGTTGAAGCGCCGCTTCAAAGTGAAGCATTTGAGA 4832
Db 4741 CCTCGAAATGCTTCCGCGCTTGGAGTTGAAGCGCCGCTTCAAAGTGAAGCATTTGAGA 4800
Qy 4833 AGCTGCTCTCTCAACCCAAAGTCCAAAAGAGCATGCTGATGAGCCCTCAGTCCAA 4892
Db 4801 AGCTGCTCTCTCAACCCAAAGTCCAAAAGAGCATGCTGATGAGCCCTCAGTCCAA 4860
Qy 4893 AGCTTAGCCCAAGGAGAGCCAGTCTGAGAAACCTTAACCAAGAAATCCCAAGACACC 4952
Db 4861 AGCTTAGCCCAAGGAGAGCCAGTCTGAGAAACCTTAACCAAGAAATCCCAAGACACC 4920

```

```

Qy 4953 AACGACCCAAAGCTTAACCGAGAGAGGCGCATCCCCAAAAGTTGAGTTCAGAGAAAGAA 5012
Db 4921 ACCGACCCAAAGCTTAACCGAGAGAGGCGCATCCCCAAAAGTTGAGTTCAGAGAAAGAA 4980
Qy 5013 ATGTCTACCAAGTCACTCAAGAGAAACAAGGCCAGGACAGAGAGACCTGACAGAGCCA 5072
Db 4981 ATGTCTACCAAGTCACTCAAGAGAAACAAGGCCAGGACAGAGAGACCTGACAGAGCCA 5040
Qy 5073 AAGGAGACCTGGCAGAAATCCTTAAGATGTTAGTTCATTTGATCAATCTGTAAGACAGA 5132
Db 5041 AAGGAGACCTGGCAGAAATCCTCGATGTTAGTTCATTTGATCAATCTGTAAGACAGA 5100
Qy 5133 ATGTGAAAACCAAGTCACTCAAGAGAAACAAGATGCTGTGTGGACCTTGAGACCAAGTTTCAG 5192
Db 5101 ATGTGAAAACCAAGTCACTCAAGAGAAACAAGATGCTGTGTGGACCTTGAGACCAAGTTTCAG 5160
Qy 5193 AGCCCATGAGATCCAGAGAGAGAGGCGGTCCTCAATGATTTTC 5232
Db 5161 AGCCCATGAGATCCAGAGAGAGAGGCGGTCCTCAATGATTTTC 5200

RESULT 2
AAV02301
ID AAV02301 standard; cDNA; 5074 BP.
XX AC
XX AAV02301;
XX AC
XX 20-JUL-1998 (first entry)
XX DT
XX DE
XX Rat tumour suppressor gene SSeCKS (truncated form).
XX KW
XX SSeCKS; tumour suppressor gene; rat; protein kinase C; mitosis;
XX KW
XX cancer; malignancy; cell proliferation; Alzheimer's disease;
XX KW
XX therapy; ss.
XX OS
XX Rattus sp.
XX FT
XX Key
XX Location/Qualifiers
XX 176..4216
XX FT
XX CDS
XX /tag= a
XX /transl_except= (pos:389..391, aa:Glu)
XX FT
XX /transl_except= (pos:1661..1663, aa:Lys)
XX FT
XX /transl_except= (pos:1668..1670, aa:Lys)
XX FT
XX /transl_except= (pos:2444..2446, aa:Arg)
XX FT
XX /transl_except= (pos:2528..2530, aa:Glu)
XX FT
XX polyA_signal
XX /tag= b
XX FT
XX polyA_signal
XX /tag= c
XX FT
XX PN
XX MO9740059-A1.
XX PD
XX 30-OCT-1997.
XX PE
XX 18-APR-1997; 97MO-US06830.
XX PR
XX 18-JUN-1996; 96US-0665401.
XX PR
XX 19-APR-1996; 96US-0635121.
XX PA
XX (GELM/) GELMAN I.
XX PA
XX (JAKE/) JAKEN S.
XX PI
XX Gelman I, Jaken S;
XX WP1; 1997-535770/49.
XX DR
XX P-PSDB; AAM31346.
XX PT
XX Tumour suppressor gene SSeCKS - used as a mitotic regulator, and
XX PT
XX inhibitor of malignant phenotype
XX PS
XX Example 6; Fig 3A-H; 162pp; English.
XX

```


Oy	2806	AACTCTCGTGTCCACCCACACGCTTCACGAGGGTTAAATGCCATATACGACTGTGCA	2865
Db	1788	AAATCTCCGTGTCTCACCCACACGCTTCACGAGGGGTCAAAATGCCATATACGACTGTGCA	1847
Oy	2866	CCAGTGAAGTGGATTTTTCACCTCAGAAAGCTGTGACACGCCACAGAGACTTCAGAGGCTCTCC	2925
Db	1848	CCAGTGAAGTGGATTTTTCACCTCAGAAAGCTGTGACACGCCACAGAGACTTCAGAGGCTCTCC	1907
Oy	2926	GTACTGAAGAAGTTACCGAAGCATTCGGGGGCCGAAGACACACACATGTGTTCGGAC	2985
Db	1908	GTACTGAAGAAGTTACCGAAGCATTCGGGGGCCGAAGACACACACATGTGTTCGGAC	1967
Oy	2986	TTTCCCAAGCTACTGACTCTCCCGACACACACAGAGAACCCACCTCAGAGGTAG	3045
Db	1968	TTTCCCAAGCTACTGACTCTCCCGACACACACAGAGAACCCACCTCAGAGGTAG	2027
Oy	3046	AGAAGTGCTGTCTAGATACAGAAAGAGAGGCGCCAGACGAGGCCATCTCCAAAGCG	3105
Db	2028	AGAAGTGCTGTCTAGATACAGAAAGAGAGGCGCCAGACGAGGCCATCTCCAAAGCG	2087
Oy	3106	TTGCAGACAAAGTGAAGAGAGATCCCAAGTGGCTTCAAACCCAGACTGTGCAGAAACGG	3165
Db	2088	TTGCAGACAAAGTGAAGAGAGATCCCAAGTGGCTTCAAACCCAGACTGTGCAGAAACGG	2147
Oy	3166	GGTCAAAGACTGTGAGAAAGGTTGAGAGGTGAAGAGACTCCGAAGTGTGGCTTCGG	3225
Db	2148	GGTCAAAGACTGTGAGAAAGGTTGAGAGGTGAAGAGACTCCGAAGTGTGGCTTCGG	2207
Oy	3226	AGAAAGAGAGAGCGTTATGCGCGAAGAGACCCGTGCAGAGAGCTGGAAGCTTGACCTTG	3285
Db	2208	AGAAAGAGAGAGCGTTATGCGCGAAGAGACCCGTGCAGAGAGCTGGAAGCTTGACCTTG	2267
Oy	3286	CACAGGCGTCTGAGACACTGAGACAGGCTACTCAGAGAGCCTTGAAAGTCTCTGAAGTCAAG	3345
Db	2268	CACAGGCGTCTGAGACACTGAGACAGGCTACTCAGAGAGCCTTGAAAGTCTCTGAAGTCAAG	2327
Oy	3346	CAGATGTAGACCATGTGCGCCACGTGCCAGGTTATCAAGCTTCAGACAGCTGATGAAACGG	3405
Db	2328	CAGATGTAGACCATGTGCGCCACGTGCCAGGTTATCAAGCTTCAGACAGCTGATGAAACGG	2387
Oy	3406	CCGTGGCCCCGTGATCATTCGAAACCTTGACACACAGTGAACAAATGGAAACACTCCCT	3465
Db	2388	CCGTGGCCCCGTGATCATTCGAAACCTTGACACACAGTGAACAAATGGAAACACTCCCT	2447
Oy	3466	TAGCAGATTTCAGACACTGTCAGATGGGACACAGCAAGATGAAACCATTTGACAGCCAGACA	3525
Db	2448	TAGCAGATTTCAGACACTGTCAGATGGGACACAGCAAGATGAAACCATTTGACAGCCAGACA	2507
Oy	3526	GTAAGGCCACTGCACTGTGTCAAGCATGCACACAGTCAACAAAGAGGCGCTTACTGCTC	3585
Db	2508	GTAAGGCCACTGCACTGTGTCAAGCATGCACACAGTCAACAAAGAGGCGCTTACTGCTC	2567
Oy	3586	AGAAAGAGAGACCTTTCGACACTTACTTATGTCTCAGGCCGAGAAAGCAATGGGAAAG	3645
Db	2568	AGAAAGAGAGACCTTTCGACACTTACTTATGTCTCAGGCCGAGAAAGCAATGGGAAAG	2627
Oy	3646	AAACGAGAAAGATGTTCTTGAACTTACACAGCAAGACTTACTGCTGCAAGCGTGGCCG	3705
Db	2628	AAACGAGAAAGATGTTCTTGAACTTACACAGCAAGACTTACTGCTGCAAGCGTGGCCG	2687
Oy	3706	TTTGTGGCAAAACGTGAGGTGGGTCAAGAGGTGAGGTGACTGTGTGATGTGAGAAAG	3765
Db	2688	TTTGTGGCAAAACGTGAGGTGGGTCAAGAGGTGAGGTGACTGTGTGATGTGAGAAAG	2747
Oy	3766	TCAAAAGAAACAGAGAGGTGTTTGTACACTCTGGAACCAACACTCAAAGAGGCTGCTGATG	3825
Db	2748	TCAAAAGAAACAGAGAGGTGTTTGTACACTCTGGAACCAACACTCAAAGAGGCTGCTGATG	2807
Oy	3826	TGACATATGACAGTGAAGTATGGGAATGGCCGGGTGTCAAGAAAGAGAGATACTGAAG	3885
Db	2808	TGACATATGACAGTGAAGTATGGGAATGGCCGGGTGTCAAGAAAGAGAGATACTGAAG	2867

QY	3866	TGAGACCTTTACGCTCGAGAGAGGAGAGATATGGAACTGACGTTGAAAAGAGAAAAGG	3965
Db	2868	TGCGAGATCTTTACCTCGAGAGAGGAGATATGAACTGACGTTGAAAAGAGAAAAGG	2927
QY	3946	AGACAAAGCCAGAGCAAGTGAAGTGAAGAGTGACACAGAAAACAGCGCTCTAGCATG	4005
Db	2928	AGACAAAGCCAGAGCAAGTGAAGTGAAGAGTGACACAGAAAACAGCGCTCTAGCATG	2997
QY	4006	AAGGAACTTACGAGAGCCAGATCTTGA CATGCGACGCTCAGAGAGGAGAGG	4065
Db	2988	AAAGAACTACGAGAGCCAGATCTTGA CATGCGACGCTCAGAGAGGAGAGG	3047
QY	4066	CACGTGGAAAGCTTGGAGAGAGCCCTTCTCCGACCAAGACAAAGACAGTTGCAATG	4125
Db	3048	CACGTGGAAAGCTTGGAGAGAGCCCTTCTCCGACCAAGACAAAGACAGTTGCAATG	3107
QY	4126	AGGTTCAAGTTCAAAGGCTTGGACACACACAGTACCTCAACACAGCAAGCTGTGAAAAAG	4185
Db	3108	AGGTTCAAGTTCAAAGGCTTGGACACACACAGTACCTCAACACAGCAAGCTGTGAAAAAG	3167
QY	4186	TCATAGAAACGGTTGTGATTTTCAGACACAGGTGAAGTCCAGAGTGTGTAGTGCACAT	4245
Db	3168	TCATAGAAACGGTTGTGATTTTCAGACACAGGTGAAGTCCAGAGTGTGTAGTGCACAT	3227
QY	4246	TATTTACAGCTGAGAGAGTCTCTGCAACGGGTGGCCATGGACTTTAGCATGACAGG	4305
Db	3228	TATTTACAGCTGAGAGAGTCTCTGCAACGGGTGGCCATGGACTTTAGCATGACAGG	3287
QY	4306	ACACGGTACCCCTGGGGCCGAGTCTCAGGACGAATTCATCCCATATATATTAATCTCTG	4365
Db	3288	ACACGGTACCCCTGGGGCCGAGTCTCAGGACGAATTCATCCCATATATATTAATCTCTG	3347
QY	4366	CTCCTGAAAGCACCTTACATCTGACCTTACAGAGAAATTAAGCGCATCCAGAGAGAGC	4425
Db	3348	CTCCTGAAAGCACCTTACATCTGACCTTACAGAGAAATTAAGCGCATCCAGAGAGAGC	3407
QY	4426	GATCAGAGAGAAAGAGACACAGCCAGATGCTGTGTCGATGCTGACGGCAAGAGAGTACG	4485
Db	3408	GATCAGAGAGAAAGAGACACAGCCAGATGCTGTGTCGATGCTGACGGCAAGAGAGTACG	3467
QY	4486	CAATCGAAAAAGCTCTCAAGGCTGAACCTGAGATCTGGAACTTGAGAGTAAAGCAACA	4545
Db	3468	CAATCGAAAAAGCTCTCAAGGCTGAACCTGAGATCTGGAACTTGAGAGTAAAGCAACA	3527
QY	4546	AGATTGTGCTGAAGCTCATTCAGACAGCCGTTGACCAAGTTCCACAGTACAGAAACAGCC	4605
Db	3528	AGATTGTGCTGAAGCTCATTCAGACAGCCGTTGACCAAGTTCCACAGTACAGAAACAGCC	3587
QY	4606	CCGAAACTCATGCTTATGATTCACAGACCCAGGTTCTGATGACAGGCTTGAACAGAGG	4665
Db	3588	CCGAAACTCATGCTTATGATTCACAGACCCAGGTTCTGATGATGCGTTGACAGCAGGG	3647
QY	4666	AGCCCAACAGTGTGTGGACAAAAATGAAGATGCCAAGTGAATAACCCATGCGCCGACG	4725
Db	3648	AGCCCAACAGTGTGTGGACAAAAATGAAGATGCCAAGTGAATAACCCATGCGCGACG	3707
QY	4726	CCAGAGAGACCTTGCAGTCTCTGACCGGTTCTGGAGGACATGGGCTCAGCCTCGAAAAATGCT	4785
Db	3708	CCAGAGAGACCTTGCAGTCTCTGACCGGTTCTGGAGGACATGG--CTCAGCTCGAAAAATGCT	3765
QY	4786	TGCGCGCTTGACAGTTGAAGACCGCGGTTGCAAAAGTAAAGCATTTGAGAACTGCGCTCTCA	4845
Db	3766	TGCGCGCTTGACAGTTGAAGACCGCGGTTGCAAAAGTAAAGCATTTGAGAAAGTGCCTCTCA	3825
QY	4846	ACCCAAAGATTCAAAAGAGACATGCTGTGATGAGCTTCAAGTCCAAAAGCTTAAAGCCACAGG	4905
Db	3826	ACCCAAAGAT--CAAAAAGACATGCTGTGATGAGCTTCAAGTCCAAAAGCTTAAAGCCACAGG	3884
QY	4906	CAGAGGCGCAGTGCTCTGAGAAACCTTAACCAAAAGAACTCCCAACACCAACGCAACAAAGC	4965
Db	3886	CAGAGGCG--AGTGTCTGAAACCTTAACCAAAAGAACTCCCAACACCAACGCAACAAAGC	3941
QY	4966	TAAACGAGAGAGGCGATCCCCCAAAAAGTTGAGGTCAGAGAAAGAAAATGTTACTCAAGT	5025

```
|||||
Db 3942 TAA CGAGAGG -ATGCCCCAAAGTTGAGTTCAGGAAGAAATGCTACCAAGT 4000
Qy 5026 CAGTCAAAGAGAAACAAGCCAGGACAGAGAGGACCTGTCAGGAGCCAAAGGAGACCTGG 5085
Db 4001 CAGTCAAAGAGAAACAAGCCAGGACAGAGAGGACCTGTCAGGAGCCAAAGGAGACCTGG 4060
Qy 5086 CAGATCCTAAGATGTTAGTTGCTCATTTGTACATCTGTGAAGACCAAGATGTGAACAAAG 5145
Db 4061 CAGATCCTAAGATGTTAGTTGCTCATTTGTACATCTGTGAAGACCAAGATGTGAACAAAG 4120
Qy 5146 TCACAGAACAAAGATGCTGCTGTGGACCTTGAGACCAAGATTTTCAGAGCCCATGAGATC 5205
Db 4121 TCACAGAACAAAGATGCTGCTGTGGACCTTG -GACCAAGATTTTCAGAGCCCATGAGATC 4179
Qy 5206 CAGAGACAGGGCGCTCCAATGATTTCCACCAGTAGACACCCCGACAAATTCAGAGGCT 5265
Db 4180 CAGAGACAGGGCGCTCCAATGATTTCCACCAGTAGACACCCCGACAAATTCAGAGGCT 4239
Qy 5266 TCATCGGAGCTAGACGACCTAACATTTCTCGTTTCAAGACTGCTTTGATTTGGCCCC 5325
Db 4240 TCATCGGAGCTAGACGAC----- 4259
Qy 5326 TTGATGCGTCCGTTGATTTCTAACATTTCTCGTTTCAAGACTGCTTTGATTTGGCCCC 5385
Db 4260 -----CTAACATTTCTCGTTTCAAGACTGCTTTGATTTGGCCCC 4299
Qy 5386 TTGATGCGTCCGTTGATTTTCGGATTTAAGGTCTCGGTTCTCAACCTTGAACCAATTC 5445
Db 4300 TTGATGCGTCCGTTGATTTTCGGATTTAAGGTCTCGGTTCTCAACCTTGAACCAATTC 4359
Qy 5446 GCCATACCTAGTTCACCTTCAAACTGGAGCATCTCTTTTATGATTTATATGATG 5505
Db 4360 GCCATACCTAGTTCACCTTCAAACTGGAGCATCTCTTTTATGATTTATATGATG 4419
Qy 5506 TTTATGAGTCTCTCTCTGACCTATGATATTTTCTAACGTTAAGCAGATGCT 5565
Db 4420 TTTATGAGTCTCTCTCTGACCTATGATATTTTCTAACGTTAAGCAGATGCT 4479
Qy 5566 TTTTGTATTATGCAATATATAACGGGTGTGAGCATAGCAGCTTTGAAAAGCTCCAA 5625
Db 4480 TTTTGTATTATGCAATATATAACGGGTGTGAGCATAGCAGCTTTGAAAAGCTCCAA 4539
Qy 5626 GCCTCAACTGTAACTGCAGCAAAACAGATAACATTCCTGCGCAAGAGACAAGTCTTT 5685
Db 4540 GCCTCAACTGTAACTGCAGCAAAACAGATAACATTCCTGCGCAAGAGACAAGTCTTT 4599
Qy 5686 TTAAGTTTACTGATGCTTAGATCTGGGCTTCTAGTCTCTGAAAAGTGGTTTTC 5745
Db 4600 TTAAGTTTACTGATGCTTAGATCTGGGCTTCTAGTCTCTGAAAAGTGGTTTTC 4659
Qy 5746 TATGCACAGCGAGCTCAGAAATAAAACCCCATTTTGAAACATCCAGATGTCCTCAATAT 5805
Db 4660 TATGCACAGCGAGCTCAGAAATAAAACCCCATTTTGAAACATCCAGATGTCCTCAATAT 4719
Qy 5806 TACCATGATTTTCCCGCTTTTGTGTAATCCAGTCCAGTTGGAAGAAGTCTCCTCT 5865
Db 4720 TACCATGATTTTCCCGCTTTTGTGTAATCCAGTCCAGTTGGAAGAAGTCTCCTCT 4779
Qy 5866 GTGTCAGATTAAAGCCCTCTCTTAATGATATGCAAAATGAGTGTGCTTAAGCCATGA 5925
Db 4780 GTGTCAGATTAAAGCCCTCTCTCTTAATGATATGCAAAATGAGTGTGCTTAAGCCATGA 4839
Qy 5926 GATGTTTCTTAATGCAGAGGAATCTGTGTGACGTTTTTTTGTGTTGACTCTTCTATGCT 5985
Db 4840 GATGTTTCTTAATGCAGAGGAATCTGTGTGACGTTTTTTTGTGTTGACTCTTCTATGCT 4899
Qy 5986 GGACCGAATTATATGCAGATCGAAGTGAGTCTCTGTTCTTTTACAGATGGTATTTGATAG 6045
Db 4900 GGACCGAATTATATGCAGATCGAAGTGAGTCTCTGTTCTTTTACAGATGGTATTTGATAG 4959
Qy 6046 ATACTGAGTTTCTGTGTTATATCTGTGCCCTTCTTTAAGAACAAATGTGTCATTATG 6105
|||||
```

```
Db 4960 ATACTGAGTTTCTGTGTTATATCTGTGCCCTTCTTTAAGAACAAATGTGTCATTATG 5019
Qy 6106 TTCCTTTGGATAAATTTGATTTGACAACTGATTTAAATAAACATATTTGACTAC 6160
Db 5020 TTCCTTTGGATAAATTTGATTTGACAACTGATTTAAATAAACATATTTGACTAC 5074

RESULT 3
ABN59861
ID ABN59861 standard; cDNA; 6614 BP.
XX
AC ABN59861;
XX
DT 28-JUN-2002 (first entry)
XX
DE Novel human coding sequence SEQ ID NO: 272.
XX
KW Human; antianaemic; vulnery; antiinflammatory; immunomodulator;
KW antiinfertility; cerobroprotective; cytostatic; rheumatic; gene therapy;
KW neuroprotective; antiparkinsonian; protein therapy; EST;
KW expressed sequence tag; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200222660-A2.
XX
PD 21-MAR-2002.
XX
PF 10-SEP-2001; 2001WO-US26015.
XX
PR 11-SEP-2000; 2000US-0659671.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX
DR WPI; 2002-292408/33.
DR P-PSDB; AB97448.
XX
PT An isolated polynucleotide for treating diseases associated with its
PT encoded polypeptide such as cancer and multiple sclerosis -
XX
PS Claim 1; SEQ ID NO 272; 509pp; English.
XX
CC The present invention provides the protein and coding sequences of 444
CC novel human proteins. These were isolated from expressed sequences tags
CC (ESTs). They can be used to stimulate cell growth, to regulate
CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
CC e.g. in burn treatment, to regulate the immune system e.g. to treat
CC multiple sclerosis, to regulate activin or inhibit e.g. to treat
CC infertility, to regulate haemostasis or thrombolysis e.g. to treat
CC stroke and cancer, to screen for drugs, to treat inflammatory conditions
CC e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.
CC Parkinson's disease. The present sequence is a coding sequence of the
CC invention.
XX
SQ Sequence 6614 BP; 2094 A; 1420 C; 1827 G; 1273 T; 0 other;

Query Match 35.4%; Score 2179.6; DB 24; Length 6614;
Best Local Similarity 64.4%; Pred. No. 0;
Matches 4223; Conservative 0; Mismatches 1794; Indels 537; Gaps 43;

Qy 3 GGCCGGCGGGAGTAGAAGGCCACTGAGCCATGCGGCGAGGAGTCCACCGAGCAGCG 62
Db 162 GGCTAGGCGGGGAGAAAGTGCGGAGGAGCCATGGGCGCGGGAGTCCACCGAGCAGCG 221
Qy 63 AGCCCGAGCAGCAGCG--GCGGGAGCACACGCCGAGGAGCTGTGCTCAGTGGCCAT 119
Db 222 AGCCCGAGCAGCAGCGCGCGGAGGAGGAGCTCCACGCCGCTGAGCCCGAGCCAGCGGCG 281
Qy 120 GGCCCGCAGCTGAAGCCTCGGA-----GCAGCTGGAGACCCCGCCGCGGAC 170
|||||
```

Db 282 GGCCCTTGCCGAGGCGGCGCCAGACACACCGGGGACCCCGCATGCTCTCGGAC 341
Qy 171 CCGCCCAACCAAGCTTCCACAGAAAGATGCGAGCTGTCTTCTGTCACGCGGTAGCTGAA 230
Db 342 CCGCGCAACCAAGCTTCTACAGAAAGATGCTGAGCTGTCCACATCAATGCGTAGCTAG 401
Qy 231 CAAGAGATGTCATGTCCTCAAGAGAAAACGAGAGG----- 268
Db 402 CAAGATAGCTCAGGCTTCAGAGAGGTGACCTAAATGCGCAAGAAAGACCTGAACGCT 461
Qy 269 -----GCAGAGAGAAAGATGCTTGAATGAGATGTTGACAGCGA 308
Db 462 CAAGAGCCCTTAACAGCCAGAGAGAAAGATGCTTGTTCACAGAGTTGAGACAGAA 521
Qy 309 GAGTCAGAAATGTGAGAGAAAAGAACCGAGTTGAAGAAATGCGCGCACTCCACAGCT 368
Db 522 GACTTCGAAAGATGTAGCGAAAGAGACTCCGATAAAGATGCTACTAAGTCAACGGTT 581
Qy 369 GTTGAAGATTCACAAAGATGCGAGAGAGAGATCAAGAAATTAATGAACAGATCCCT 428
Db 582 GTTCAAGCATCAAGATGATGGCAGAGAGACACCGAAATATGAAACAGATTCCT 641
Qy 429 GCTTCAGAAAATATGTGAGAGAAATGATACAGCTGCTGATGCCAGGCTATGATGTT 488
Db 642 TCTTCAGAAAGCAATTTAGAGAGCTTAACACACCACTGAGTCCAGGCTTAATGATATT 701
Qy 489 GGCTTCAGAAAATATTTAAATTTGTGTTTAAATTCACGCGTGAAGAGATTAAT 548
Db 702 GGATTTAAGAAAGTGTTTAAGTTTGTGCTTTAAATTCATGTAAGAAAGATTAAGACA 761
Qy 549 GAAAAGTCAGATCTGTCTCACTACTCTCTCAAGAGAGATGAAGGCGAAGGGCGACAA 608
Db 762 GAGAAAGCTTGACACTGTCTCACTACTCTGTGAAGAAAGATGAAGGGAGGAGACAGCA 821
Qy 609 GCCTCTGTCGAGCTGAGAACCCAGAGAGCCAGTGTGAAGTGCCTCGAGAGTCA 668
Db 822 -----GGGGCTGGCGACCAACAGAGACCCAGCTTGGGGCT-----GGAAGAGA 866
Qy 669 GCATCCAAAGAAATGAGCTGAAGCAATCAGAGAGACAAAGAGACCTTGAAGCAA 728
Db 867 GCATCCAAAGAAAGCAAGCAACCAACCACTAAGAGAAACCGAAGAGACCTTGAAGCAAT 926
Qy 729 GAACGAGACAGACAGAAATCCCTCTTCAAGCGCAATCTGATCAAGCGCTGAGAGAA 788
Db 927 GAGCAAGCCAGCAGCAAAATTTCTCCCGAGCGAATCTGGCCAGACA---GTGAGGAA 983
Qy 789 GCCAAGATGAAGAGAGAAAGAAACAGAGAAAGAGCCACCAAGTCCCAATCCCG 848
Db 984 TGCAAAG 1043
Qy 849 AGCAGCCCAAGTCAGACAGTGAACAATCTTCTTCAAGAGTTCTTCACTACGCTTG 908
Db 1044 ACTAGTCCCGTACCAAGTGAAGAACAGATCAACCTTCAAAAAATTTCTTCACTCAAGGTTG 1103
Qy 909 GCGCGCTGCGCAGAGAAACAGCTTCAAGAAATCAAAAGAGATGATCTGAGAAATGCGC 968
Db 1104 GCGCGCTGCGCAGAAAGAGCAAGTTTCAAGAAAGCCGAAGAGATGAAGTGAAGTTTA 1163
Qy 969 GAGAGAGAGAAAG 1028
Db 1164 GAG 1223
Qy 1029 CCAGCTTCGAGAG-----GCAGAGCCGCGAGAGAGC 1061
Db 1224 GTTTCCTCGAGAGAAATGACCGGCTCCGAGCAAGCCCAACCAAGAGCGGAGAGAAAT 1283
Qy 1062 ACAGACCAAGCCAGGTTGTACAGACATACAGAGAGTGAAGTGTCTTTGAGAGACAG 1121
Db 1284 GCCCAG 1343
Qy 1122 GTTGTGAGCTGAG 1181
Db 1344 GTGAGTGGCTCGAGAGGACCTTCTGAAGAGAACTGTCTCGTTGGCGACAGAGTGT 1403

Qy 1182 GATGAGAAAGATGGAAGCCCAAGAA---GTTTTCAGAGGTCAGAGTGAAGCAGCTG 1238
Db 1404 GATGAGAAAGATGGAAGTCCACCAAGAAAGATGTGGCCAGAGTCAAGTCAAGACCTG 1463
Qy 1239 GAGAGAGACAG 1298
Db 1464 GAG 1499
Qy 1299 GGAACAGAGAAATCTTGGCCCTGAGAAATCTGCTGAGCCCAAGAGATCCCCAGAA 1358
Db 1500 GAAACAGAGAGATGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1559
Qy 1359 GCTAGCTGCTGAG 1418
Db 1560 GCCAAGCTTCGAG 1619
Qy 1419 ACTCAACTGACAGACCTAAGTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1478
Db 1620 ACACAG 1679
Qy 1479 GTCAAG 1538
Db 1680 GTGAG 1739
Qy 1539 AAGAAATCTTCAAGTCAAGTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1598
Db 1740 AAGAAATCTTTCACAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1799
Qy 1599 CGAG 1658
Db 1800 AGAGAG---GAG 1856
Qy 1659 GAGAGTCTGATGAGCAG 1718
Db 1857 GACAGCCAG 1916
Qy 1719 ACCAGTGTCTGAG 1778
Db 1917 ATCAGTGTCTGAG 1976
Qy 1779 ACTATTCGATGAG 1838
Db 1977 GCTATTCGATGAG 2036
Qy 1839 ATGTGACACCCAG 1898
Db 2037 ATGTGACGCGCCAG 2096
Qy 1899 GAG 1958
Db 2097 GACAAAGTCAAG 2156
Qy 1959 GAAGTCAAAATCTGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2018
Db 2157 GAAATGAAG 2216
Qy 2019 TCAGTGTCTTGGAG 2078
Db 2217 TCAGTATCTTGGAG 2276
Qy 2079 TCCTCTTCAAGATGAG 2138
Db 2277 TCCTCTTCTGATGAG 2336
Qy 2139 GAGGCAG 2198
Db 2337 GAGGCAG 2396
Qy 2199 GACCAAGGCAAG 2258
Db 2397 GATCCAGGAG 2456

QY 2259 GTCTCACTTGGAGTCAATTTAAAGATTAGTCACTCCAGABAAAAATCAAGTCAAAA 2318
DB 2457 GTTTCACCTCGGAGTCAATTTAAAGGTTAGTACGCCCAAGAAAAAATCAAGTCCAAG 2516
QY 2319 CTGGAAGAGAAAGCCGAAGAC-----TCTAGTGTAGAGCAAGTTGTCCACTGAG 2366
DB 2517 CTGGAAGAGAAAGCCGAAGACTCCATAGCTGGGTCTGGTGTAGAACAATCCACTCCAGAC 2576
QY 2367 ATCGAAACCGGTAGAGAAATCTTTGGGTTTCCATTAGAAATTCATCCCGGACGGGG 2426
DB 2577 ACTGAACCCGGTAAAGAAATCTGGGTCTCAATCAAGAAATTTATTTCTCGACGAAGG 2636
QY 2427 AAGAAAAAGGACAGCGGAAGCAAGCAAGCACTCTGGAAGACTCAGGCGCAGTGGAG 2486
DB 2637 AAGAAAAAGGACAGTGGAAACAAGAACAGCCCTGTGGAAGACGAGGGCCCAACAGGG 2696
QY 2487 ATAAATGAGGACGACCTTAATGTCCAGCGCTCGTCTGTCTGAGTATAATGCAAGTG 2546
DB 2697 GCCAACGAATGACTCTGATGTCCTCGGCGGTGCTCTGTCTGAGTATGATGCTGTA 2756
QY 2547 GAGAGGAGAAGATGG-----AAGCCAGGGGAATACGGAGCTGCCCAAGCTGCTG 2597
DB 2757 GAAAGGAGAAAAATGGAGGACACAGAACGCCAAAAAGCGCAGAGCAGCCGACGAAG 2816
QY 2598 GGGCTGTGTACGTCTCGAGGAGCTCAGTAAGACTCTGGTCCACACTGTGAGTGTCCGA 2857
DB 2817 GCAGCCACTGAGGTCTCAAGAGCTCAGCGAGAGTCAAGTTTCAATATGATGGCAGAGCT 2876
QY 2658 GTCAATTGATGGACAGGGCAGTCAACAGTGTGAAAGCGGTCTCTTGTGATATCC 2717
DB 2877 GTCGTACGGGACAGGGCAGCTACCAATATTGAAAGAGTCTCTCTTTGATATCT 2936
QY 2718 GCTTCGTAAACAGAACTCTTTGAAACACACAGCGGGAGAGCATGCCACCTGTTTGAAGAG 2777
DB 2937 GCTTCAGTGCAGAACTCTTTGAACAAGTAGAAGCTGAAGCCGCACTGTTAACTGAGGAG 2996
QY 2778 GTCACTGAAAAGACATCATTCG---AGAGAAACTCTGTGTCTACCCAGACGTTACCA 2834
DB 2997 GTATTGAAAAGAGAAGTAATTTGCAAGAGAAGAACCCCCACGGTTACTGAACTCTGCCA 3056
QY 2835 GAGGTTAAAGATGCCATGACGACATGGTCAACAGTGAAGTGAATTTCACTCAGAAAGCT 2894
DB 3057 GAGAACAGAGAGGCCCGGGGGACACGGTCTGTTAGTGAAGCGGAATTTGACCCCCGAAGCT 3116
QY 2895 GTGACAGCCACAGAGACCTCAGAGGCTCTCCGTACTGAAGAAGTTTACCGAAGCATCCGGG 2954
DB 3117 GTGACAGCTGCAGAAACTGCAGGGCCATTTGGTTCCGAAGAAGAAACCGAAGCATCTGCT 3176
QY 2955 GCCGAAGAGACCAAGACATGGTGTCCGAGTTTCCAGCTGACTGACTCCCGACACCC 3014
DB 3177 GCTGAAAGAGACCAAGAAATGGTGTACAGCAGTCTCCAGTTTAAACCGACTCCCGACACCC 3236
QY 3015 ACAGAGGAAGCCACCCAGTTTCAGAGGTAGAGTGTGTGTAGATACAGAAAGAGAG 3074
DB 3237 ACAGAGAGGCCATCTCCGTGCAGAGGTGAAGTGGCGTACTGTACATAGAAGACAA 3296
QY 3075 GAGCCGACAGCGCAGGGCCATCTCTCAAGCCGTGTGCACAGAAGTGAAGAAGAGTCCCGAG 3134
DB 3297 GAGAGGCGGACTCAAGAGGTCTCCAGCAGTGCACAGAAAAGTGAAGAAGGAATCCAG 3356
QY 3135 GTGCTCTGAACC---CAGACTGTGCAGAGAACCGGGGTCAAAAGCACTGGAAGAGTTGAG 3191
DB 3357 CTGCTTGGCACCGGTGGGCGCAGAAAGATGTGCTTCAGGCTGTGCAGAGAGCAGAGGCAGAA 3416
QY 3192 GAGGTAGAGGAGGACTCCGAAGTCTCGCTTCGGAGAAAGAGAGGAGGAGTATGCCAAA 3251
DB 3417 AGACCAAGAGAGCGGCTGAAGCGTCCGGTCTGAAGAAAGAGAGCGGATGTAGTGTGAAA 3476
QY 3252 GGACCCGTGCAGGAAGCTGGAGTCTGAGCATCTTTGCACAGGGCTCTGAGACTTGGACAGGCT 3311
DB 3477 GTAGATGCTCAGGAGGCAAAAACCTGAGCCCTTTTACACAGGGAAGGTGGTGGGACACCC 3536
QY 3312 ACTCCAGAGAGCCCTTG---AAGTTCTGAAAGTCAACGCGCAGATGTAGAC-----CAT 3359

DB 3537 ACCCCAGAAAAGCTTTTGA AAAAGCTCTCTCAAGTCAAGAGAGCATAGAGTCCAGTGAGCTT 3596
QY 3360 GTGCCACAGTGGCA-----GGTTATCAAGCTCCAGCAGAGCTGATGGAA 3401
DB 3597 GTAACCACTTGTCAAGCCGAAACCTTAGCTGGGTAAATCAGAGGATGGTATGGAA 3656
QY 3402 CAGGCGGTGGCCCTCGATGATCCGAAACCTTGACAGACAGTGAAGCAAAATGGAAGCACT 3461
DB 3657 CAGGCTATCCCCCTGACTCGGTGGAAACCCCTACACACAGTGAAGTGTGGAAGCACC 3716
QY 3462 CCCTTAGCAGATTACAGACACTGAGATGGGACACACAGACATGAAACCATTTGACAGCCAG 3521
DB 3717 CCCGTAGCCACTTTGACGCAACCCAGGCAACCCAGAAAGCAGAGATTGTGGAATTCAT 3776
QY 3522 GACAGTAAAGCCACTGTCAGCTGTGAGCAGTCAAGGTCAAGAGAAAGAGGGCGGTACT 3581
DB 3777 GAGGAGNATGAGTGCATCTGGTACCCAGTCAAGGGGGACAGAGCAGAGGAGTCTCT 3836
QY 3582 GCTCAGAAAGAGAGGCTTCGACACTACCTAATAATTTGTTCCAGCCAGGAAGAACATGGG 3641
DB 3837 GCACAGAAAGAGAGGCTTCAGCACCTTCAGTCTTCCAGTTTGTGTTCCAGGAAGAACTAAAGAA 3896
QY 3642 GAAGAACCCAGGAAGAGATGTTCTTGAACCTACACAGCAAGAGCTTACTCTGACGCGTG 3701
DB 3897 CAATCAAAAGATGGAAGACACTCTAGAGCATACAGATAAAGAGGTGTCAAGTGAATCTGTA 3956
QY 3702 CCCGTTCTGCAAAAGACTGAGGTGGGTCAAGAGGGTGA-----GGTTGACTGGTTG 3752
DB 3957 TCCATTCTGTCAAAAGACTGAGGGGACTCAAGAGGCTGACCAAGTATGCTGATGAAAAACC 4016
QY 3753 GATGGAGAAAAAGTCAAAAGAAACAGGAGGTGTTTGTACACTCTGG-----ACCCAAC 3806
DB 4017 AAAGACGTACCAATTTTTCGAAGGACTTGAGGGGTCTATAGACACAGGCATAACAGTCACT 4076
QY 3807 AGTCAAAAGCTCTCATGTGACATATGACAGTGAAGTGAAGTGGAGTGGCGGGGTGTCAAG 3866
DB 4077 CGGGAAGAGGTCTCACTGAAGTGGCTTAAAGGTGAAGGGACAGAAAGCTGAATGTAAA 4136
QY 3867 GAAAAGCAGAGTACTGAAGTGCAGAG-----TCTTAGCCTGGAGGGGA 3911
DB 4137 AAGGATGATGCTCTTGAACCTGCAGAGTCAAGTCTCTCCATCCCTCCCGGTGGAGAGA 4196
QY 3912 GAGATGGAAAACCTGACCTTGAAAAGGAGAAAGGAGACAAAGCCAGAGCAAGTGAAGTGA 3971
DB 4197 GAGATGTTAGTTCAAGTTCGAAAGGGAGAAAAACAGAAAGCAGAGCCAAACCCATGTGAATGAA 4256
QY 3972 GAAGG---TGAGCAGGAAACAGCCGCTCTGAGCATGAAGGAACCTACGGGAAGCCAGTC 4028
DB 4257 GAGAAGCTTGAGCACGAAACAGCTGTTACCGTATCTCAAGAGGTCAAGTGAAGCAGCTCCTC 4316
QY 4029 CTCACACTTGCATGCCAGCTCAGAGAGGGGAGGAGGCACTGGGAAGCCTTGGAGGAAGC 4088
DB 4317 CAGACAGTGAATGTCCTCCATCATAGATGGGGCAAGGAAGTCAAGTCTGGAAGGAAGC 4376
QY 4089 CTTTCTC---TCCACAGCAACAGCAAGAGGTGTGATAGAGGTTCAGAGTTCAGAGCCTG 4145
DB 4377 CCTCTCTCCCTGCTAGGTCAAGAGGAGGAGTATGACCAAAATTCAGATTTCAGAGCTCT 4436
QY 4146 GACAAACAGTCTCAAAACAGCAGAAAGCTGTGGA AAA-----GGTCATAGAAAACGGTT 4199
DB 4437 GAGGCATCATTTCACTCTAAACAGCGGTGAGAGGAGGAAAGGTCTTAGGAGAAACTGCC 4496
QY 4200 GTGATTTTCAGAGACAGTGAAGTCCAGAGTGTGAGGTGCACACTTATTACAGCTGAG 4259
DB 4497 AACATTTTAGAAAACAGGTGAACGTTGGAGCCTTCAGAGGTGCACATTTAGTTCTTGAAGAG 4556
QY 4260 AAGTCTCTGCAACCGGTGGCCACTGGACTCTTCAGCATGCAGAGACACGGTACCCCTG 4319
DB 4557 AATCTCTGAAAAAATGAAGACTTTGCGCTCATCCAGGGGAAGATGCTGTGCCACA 4616
QY 4320 GGGCTGAGTCTCAGCGAGAAATCCATCCCAATCATAGTAACCTCTCTCTCTGAAAGCACC 4379

4617 GGGCCCACTGTCAGGCAAAATCCAGACACGATGATGATCTGCTACTTACCAAGAAAGGC 4676
4380 CTACATCTCTGACTTACAGAGAGAAATTAAGCCATCCAGAGAGAGCCATCAGAGAGAG 4439
4677 TTAAGTTCGACCTGAGAGAGAGAAACCAATCATCTGAATGGAGATGAGATGATGATC 4736
4440 GACAAAGCCAGATGCTGCTGCTGATGCTGACGAGAGAGAGATGACGATGAGAAATGTC 4499
4737 GATGAGCGAGTGTGCTTGGC-----AGGAGGTCAAGTGAATGATGAGAGAGAT 4790
4500 CTCAAGGCTGAACCTGAGATCTGAGACTTGAAGTAAGCAACAATTTGCTGTAAC 4559
4791 TTAGAGCTGAAATGGAATTTTGAAGCTTGAAGCAAAAGAGATGATCTTCCAAAC 4850
4560 GTCAATTCAGAGAGCCGTTGACCACTGTCACAGTAC---AGAAACAGCCCGGAACTCAT 4616
4881 ATCATCCAGACAGCCGTTGACAGTTGTGATGCTAAGAAAGAACGCCAGAAATGTTG 4910
4617 GCTTATGATTCACAGACCAAGTTC-----CTGANTCAGAGCTTGACAG 4660
4911 ACCTCTGATTTACAGACCAAGCTCAGTATMAAGCTGACAGCCAGAGACGCTGACAG 4970
4661 CAGGAGGCCCAACAGATGCTGGAACAAATGAAGATGCAAGATGAAACACCCAGTGC 4720
4971 GAACCGAGAAAGAGAGAGAGAGACCTCAGGCTCTGACAGAGATGAAACACCAATTA 5030
4721 GCAGCCCA-----GAGAGACTTGAAGTCTGACCGTCTGGA 4759
5031 TCAGCCAAAGAGAGATCAGAGTCAACCGAGTGGGCAAGACATTTCTGATTTTCCAAA 5090
4760 GGCATGAGCTCAGCTCGGAAATGCTGCTGCGGCTTGACAGTTGAAAGCCCGGTGCA- 4818
5091 GACATGAGTGAAGCTCAGAAAGAACATGACTGTTGAGTGAAGTTCACTGTAAT 5150
4819 ----- 4818
5151 GATCAGCAGCTGGAAGAGTCTCTCCCATCTGAGAAAGAGAGGTGAGACTGGAACA 5210
4819 -----AGTAAAGATTGAGAGCTGCT 4840
5211 AAGTCTGTCAGAGATGATGTCATGCTTGTAGAGAAAGATAGAGAGTACATTA 5270
4841 CCTCAACCAAGAA--TCCAAAGAGACATGCTGCTGATGAGCCCTCAGCTCCAAAGCTTA 4898
5271 GTTGAACCGAAAGAGATGAAAGAGTGAATGATGATGATGATGATGATGATGATGATG 5330
4899 GCGCAGGAG 4958
5331 GCGCTGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5390
4959 CCAAGCTTAAACGAG 5018
5391 CCAAAACAAAG 5450
5019 ACCAAGTCACTAAAG 5078
5451 AGTGAATCAGATTAAG 5510
5079 GACCTGGAGAGATCTTAAGATGTTAGTGS-----CTCATTTGATC 5117
5511 GATCTGCAAGTCAAGATCTTACAGATCTTAAACATATGATGATGATGATGATGATGATG 5570
5118 ATCTGTAAGACAGATGTAAGAAACAGTCAAGAAACA-----GATGCTGCTTTGGAGC 5173
5571 GTTTGGAAGACAGATGTAAGAAACAGTCAAGAAACA-----GATGCTGCTTTGGAGC 5629
5174 CTTGAGACCAAGATTTCAAGAGCCATGATGATCAGAGAGAGAGAGAGAGAGAGAGAGAG 5233
5630 CTGAAGACCAAGATTTCAAG-ACCTTGAGATTTGAGAGAGAGAGAGAGAGAGAGAGAG 5688
5234 ACCAGTGAAGACACCCGCAATTTGAGAGCTTCACTGAGAGAGAGAGAGAGAGAGAGAG 5293
5689 TTTTGAAGAGC-CCCTGAACATCTGAGGCTTCACTGAGAGAGAGAGAGAGAGAGAGAG 5747

5294 TCCCTGTTCAAGACCTGCTTGAATTTGCCCTTGATGAGCCGCTCCGTATTTCTAACATT 5353
5748 TCCCTTTCCAAAGACCAACTACATATTTCCCTTGATTAAC----- 5787
5354 TCCCTGTTCAAGACCTGCTTGAATTTGCCCTTGATGAGCCGCTCCGTATTTCTAACATT 5413
5788 -----CATTAATTTCTGATTTA 5805
5414 AGTCTGCGCTCTCAACTGGAAC-CAATTTGCGCACTTACCTTACCTTCAACT 5472
5806 AGTCTTAAATTTCTTAACTGGAACCTGAGTTGGCAATTAATCTGCTTCTGGAAC 5865
5473 GGAGATCTCTCTTATGATTTATATGATGTTTATATGATGTTTATGATGTTTATGATGTT 5532
5866 GGAGATCTCTCTTATGATTTATATGATGTTTATATGATGTTTATGATGTTTATGATGTT 5592
5533 TGTATATTTTCTTAAAGTT-----AAGCATGCTTTTGTATTA 5575
5923 TGTATATTTTCTTATGTTTAAAGAAATGTCAGAGATATCAATGCTTTTGTATCA 5982
5576 TGCAATATATTAACGAGTGTGCAAGCATAGCGAGCTTTGAAAGCTCCAAAGCTCAAG 5635
5983 CAGATATATGATGAGGAGATGCTCATATGTCAGAGCTTGGGAGCTTTAAGCTCAGTTA 6042
5636 TAACCTGAGCAACATATTAATTC-----CTGGCAAGAGAGACAAGCTTTT 5686
6043 TATTAACCAAAAG 6102
5687 TAAAGTTACTGATGCTTATGATCTGAGGCTTCTAGTCTCTGAAAGTGTGTTTCTCT 5746
6103 TAAATTCATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6182
5747 ATGCAAGAGAGCTCAGAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5806
6163 ATTAACAG 6222
5807 ACATGATTTTTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 5866
6223 ATTGGAAATTTTGC-----TTTCAACCCAGTGGAGGTTTGAAGAGATTTATTTCT 6275
5867 TGTGAGA-----TTAAGCCCTGCTCTTATATGATGATGATGATGATGATGATGATGAT 5909
6276 GGTAGCAATTTAATCTTATCATCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 6335
5910 GTGCTAAGGCGATGAGA-----TGTTCCTAATGCAAGAGAACTGT 5953
6336 GTGCTTATCTGAGGAG 6395
5954 TGTAGTTTTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6013
6396 TGTAG--TTTTTATTTCTTATATGATGATGATGATGATGATGATGATGATGATGATGAT 6453
6014 AGTCTGTTCTTATCAGATGATTTGATGATGATGATGATGATGATGATGATGATGATGAT 6073
6454 AGTCAAGTCTTTTACAAATGATTTTATGATGATGATGATGATGATGATGATGATGATGAT 6513
6074 TGCCCTTCTTAAAGACAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6133
6514 TGCCATCTTTTAAAGACAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6573
6134 CTGATTTTAAATTA 6147
6574 CTGATTTTAAATTA 6587

RESULT 4
ABN96942
ID ABN96942 standard; DNA; 6608 BP.
AC
XX ABN96942;
XX
DT 13-AUG-2002 (first entry)

Db 1497 GAAACAGAGGCTGTGCGCAGCTGAAAGATTGGTTGATGATGAGAACTCAGGAA 1556
 Qy 1359 GCTGAGCCTGTGAGAGCTGATGAAAGCAGAGATGTGTCTCTGAGAGAACAC 1418
 Db 1557 GCCGACCTGCGAAGAGCTGTGAGACTCAAGAAAGTGTGTTCCGAGAGAGACCTT 1616
 Qy 1419 ACTCACTGACAGACTTAAGTCTGAGAGAGAACGCTGCCAAACACCGAAGGACTT 1478
 Db 1617 ACACAGGAGCTGACCTCAGTCTGATGAGAAAGGTGCTCAAAACCCCGAAGGCTT 1676
 Qy 1479 GTCAGTGGTGGAGATGCTGCTCTCAGAGAAAGATCAAGGTACAGAGAGTCTCTTG 1538
 Db 1677 GTGAGTGGGAGAAATCTGTCTATCAGAGAGAAATGAAGTGCAGAGAACTCACTA 1736
 Qy 1539 AAGAACTCTTCAGTACTCAGGCTTAAAGAACTGTCTGGAGAGACAGAGGGGAAA 1598
 Db 1737 AAGAACTTTTACAGACACTGGCTTAAAAAGCTTCTGAGAAAGAAACAGAAAGGAAA 1796
 Qy 1599 CGAGAGGTGGGGAGACGAGAGCCTGAGAAATCAACACATTCACACCGAATCCCA 1658
 Db 1797 AGAGG---GGAGAGACGAGGAATCAGGGGAGCACACTCAGGTTCCAGCCGATCTCCG 1853
 Qy 1659 GAGAGTCTGATGAGACAGAGGAGAGAGCTGTGCTGCTCCCGAGAGAGCCTGAGAG 1718
 Db 1854 GACAGCCAGAGAGAGCAGAGAGGCGAGAGCTTGCCTCATCCCTGAGAGAGCCGAGAG 1913
 Qy 1719 ACCAGTGTCTGAGAGAAAGGCGCGTGGAGCACCCAGAGTGGGGAGCTGAGAGAG 1778
 Db 1914 ATCACTGTCTGAGAAAGGCTTACCGAGGTGCAAGAGATGGGGAAGCTGAAAGAG 1973
 Qy 1779 ACTACTTCGATGAGAGAGAAAGAGAGAGATCATCTCCCTGGGCACTCTTCAAAA 1838
 Db 1974 GCTACTTCGATGAGAGAGAAAGAGAGAGGTGTCACTCCCTGGGCACTCTTCAAAA 2033
 Qy 1839 AAGTGAACCCAGAGAAAGGGGTCGAGAACCTTCTGAGTGAACAAGAGAGAAAGCTG 1898
 Db 2034 ATGGTGAAGCCAGAGAGGCTGTGAGACGCTTCGAGAAAGTGAATAAGATGAGCTG 2093
 Qy 1899 GAGAGGTCAAGAGCGCCACTTGTCTCCTCAGTGAAGACAGTGTCAAGAAATGCAAG 1958
 Db 2094 GACAAAGTCAAGAGCGCTACCTTGTCTCAGCGAGACAGCCTGAAATGCAAGAA 2153
 Qy 1959 GAAGTCAAAACCTGTGTGAGAGAAACAAACCCAGAGAAACAAAGCTGAGGTGATCT 2018
 Db 2154 GAAATGAAAGGAGCGTGGAGAGCCAAAGCCGAGAAACAAAGGCGAAGTGTATCC 2213
 Qy 2019 TCAGTCTTTGGAGACATGATTTGTGTGATCATCCAGAGAGAGAGAGAGCA 2078
 Db 2214 TCAGTATCTTTGGAGCTTTAATTTGTGTGATCATCCAGAGAAAGAGAGAGAG 2273
 Qy 2079 TCTCTTTCAGATGATGAGAGAGGCGCAAGACACTGGAGGGGAGTCAAGAGCAGAG 2138
 Db 2274 TCTCTTTCATGATGAGAGAGGCGCAAAAGCAATGGAGAGAGACCAAGAAAGTCAAT 2233
 Qy 2139 GAGGCGAGCAAGACAAAGAGCGGAGACAGAGCTGTCTCTCCAGAGACCCAGAGCAG 2198
 Db 2334 GAGGCGGAGAAAGACAAAGAGCGGGAGACAGAGGATCTTCTGCTTCCAAAGACAT 2293
 Qy 2199 GACCAAGCGGAGAGAGTCTCACCAGGCGAGCGGAGAGCCCTTCCAGAGGGAGAGT 2258
 Db 2394 GATCCAGGGGAGAGAGTCTCCTCCGAGAGCTGAGAGCCCTTACCGAGAGGGAGGCG 2453
 Qy 2259 GTCTCACTTTGGAGTCAATTAAGATTAGTCACTCCAGAGAAAGAAATCAAGTCAAAA 2318
 Db 2454 GTTTCACCTGGAGATCATTTAAAGTTAGTCAAGCAGAGAAAGAAATCAAGTCAAG 2513
 Qy 2319 CTGAGAGAGAGAGCCGAGAGC-----TCTAGTGAAGCAGTGTCTCACTGAG 2266
 Db 2514 CTGAGAGAGAGAGAGAGAGCTCATAGCTGGGTCTGGTGTAGAACTTCCATCTCAGAC 2573
 Qy 2367 ATGGAACCGAGTGAAGAAATCTTGGGTTTCACTTAAGAAATTCATCCCGAGAGCGCG 2426
 Db 2574 ACTGAACCCCGTAAAGAAATCTGGGTCTCATCTCAAGAGATTATTTCTTGAAGAGAG 2633

Qy 2427 AAGAAAGGCGAGACGAGAGAGCAAGAAACAGCCACTGTGAGAGACTCAGGGCCAGTGA 2486
 Db 2634 AAGAAAGGCGAGATGGAGAAACAAAGAACAGCCCTGTGTAAGACGAGGCGCAAGAGG 2693
 Qy 2487 ATAAATGAGAGAGACCTTAATGTCTCCAGCCGTGTGCTCTGTCTGATTAATGCACTG 2546
 Db 2694 GCCAAGAGATGATCTGTATGTCTCCGCGGTGTCTCTGTCTGATGATGTCTGTA 2753
 Qy 2547 GAGAGGAGAGAGATG-----AGCCAGGGGAAATACGAGAGTCCCGAGCTGCTG 2597
 Db 2754 GAAAGGAGAGAAATGAGAGGACAGCAAGCCCAAAAGGCGAGAGAGCCGAGAGAG 2813
 Qy 2598 GGGCTGTGATGCTGTCCAGAGAGCTAGTAAGACTGTGTCAACTGTGATGTGCA 2657
 Db 2814 GCAGCACTGAGGTGTCCAAAGAGCTCAGGAGAGTCAAGTTCAATAGATGGAGAGCT 2873
 Qy 2658 GTCAATTGATGGAGCAGAGGAGTCAACCATGTGTGAAGAGCGGTCTCTTGTGATATCC 2717
 Db 2874 GTGCTGACGGGAGAGAGGAGGAGTACCATTAATGAGAAAGTCTCTTCTGTGATATCT 2933
 Qy 2718 GCTTCCTTAACAGAACTCTTGAACACAGCGGAGAGAGTCACTGATCCACTGTTGAAG 2777
 Db 2934 GCTTCAGTGAAGAACTCTTGAACAAAGTGAAGCTGAAGCCGACGTTAACCTGAGAG 2993
 Qy 2778 GTCACTGAAAAAGCATCATTTG---AGAGAACTCTGTGTCTCACCCAGAGCTTACCA 2834
 Db 2994 GTATTTGAGAAAGATGATATTTGCAAGAGAGAACCCCAAGTATCTGAACTCTGCA 3053
 Qy 2835 GAGGATTAAGATGCGCATGACAGCATGTGTCAAGTGAAGTGAATTTACTCTAGAACT 2894
 Db 3054 GAGAAAGAGAGGCGCGGGGCGACACGGTGTGTAGAGGGAGATGACCCCGAAGCT 3113
 Qy 2895 GTGACAGCCACAGAGACTCTGAGAGCTCTCCGTACTGAAGAGTATTCAGAGATCGGG 2954
 Db 3114 GTGACAGCTGAGAACTGACAGGCGCATTTGGTTCCAGAAAGAGACGAAAGCATCTGCT 3173
 Qy 2955 GCCGAGAGACCAAGACATGATGTGCGCAGTTTCCAGCTGACTGACTCCCGAGACACC 3014
 Db 3174 GTTGAAGAGACCAAGAAATGTGTGACAGTCTCCAGTTAACGACTCCCGAGACACC 3233
 Qy 3015 ACAGAGAGAGCAACCCAGTTCAAGAGGTGAAGATGTGTCTAGATTAAGAAAGAG 3074
 Db 3234 ACAGAGAGGCGACCTCGGTGAGAGGTGGAAGGTGCGTACCTGATGAAGAGCA 3293
 Qy 3075 GAGCGCCAGAGCGAGCCATCTCTAACCGTTGTGAGACAAAGTGAAGAGAGTCCAG 3134
 Db 3294 GAGAGCGGAGCTCAAGAGGTCTCCAGGCGAGTGCAGAAAAAGTGAAGAGAGATCCAG 3353
 Qy 3135 GTGCTGCAAC---CAGACTGTGAGAGAAACGGGTCAAAAGCACTGAGAGGTTGAG 3191
 Db 3354 GTGCTGGACCGGTGGGCGAGAGATGTCTTCAAGCTGTGCAAGAGAGCAGAGCAGAA 3413
 Qy 3192 GAGTGAAGAGAGACTCCAGAGTGTGTGTGAGAGAAAGAGAGAGACTTAATGCGAAA 3251
 Db 3414 AAGACAGAGAGAGAGCTGAGAGTGTGAGAGTGTGAGAGAGAGAGAGTGTGAGAA 3473
 Qy 3252 GAGCCGTGCAAGAGCTGAGAGTGTGAGATCTTGAACAGAGGCTTGAAGTGAAGAGT 3311
 Db 3474 GTAGATGCTCAGAGAGCAAAACCTGAGCTTTTGAACAAAGGAGAGTGTGAGAGAGC 3533
 Qy 3312 ACTCAGAGAGAGCTT---AAGTCTGTGAAGTCAAGGAGATGTAGC-----CAT 3359
 Db 3534 ACCCAGAGAGCTTTGAGAAAGCTCTTCAAGTCAAGAGAGATGAGTCAAGTGTGAG 3593
 Qy 3360 GTGCGACAGTGC-----GGTTATCAAGCTCAGAGAGTGTGAGAA 3401
 Db 3594 GTTACACATTTGTCAAGCGAGAACTTACGTGGGTGAATATCAAGAGAGTGTGATGAA 3653
 Qy 3402 CAGGCGGTGGCGCTGAGTATCCGAACCTTGAACAGAGAGTGAAGCAATGGAAGCACT 3461
 Db 3654 CAGGCTATCCCTGAGCTGTGAGAAACCTTACAGACAGTGAAGCTGATGAGAGCACT 3713

Db 5800 AGTCTTAATTTCTTAACCTGGAAGTGGCAATCACTAGTTCTGCTTGAAGT 5859
 Qy 5473 GGAGCATCTCTTTATGATTTATGATGTTTATGATGCTCTCTGTAACAT 5532
 Db 5860 GGAGTATCATCTTTACATATTTATATGTTTATGATGCTCTCTGTAACAT 5916
 Qy 5533 TGTATATTTTCTTCAAGTTT-----AAGCAGAGCTTTTGTATTA 5575
 Db 5917 TGTATATTTTCTTCAAGTTT-----AAGCAGAGCTTTTGTATTA 5976
 Qy 5576 TGCATATATATAGCGGATGTCAGCATGAGCTTTGAAAGCTCCAGGCTCAACTG 5635
 Db 5977 CACAGTATATATGAGGATGTCAGCATGAGCTTTGAAAGCTCCAGGCTCAACTG 6036
 Qy 5636 TAACCTGAGCAACAGATTAACATTC-----CTGGCAAGAAAGAGCAAGCTTTT 5686
 Db 6037 TATTAACCAACAAAGAGAGCTCTAGATGATTAACATTCCTGATCAAGTAACTCTT 6096
 Qy 5687 TAAAGTTACTGATGCTTATGATGCTGAGCTTCTGATGCTGAAAGGTTGTTTCT 5746
 Db 6097 TAAATTCATTAATGATGAGTCCATATTTAGTGTGTTAGTGTGTTGTTGTTCT 6156
 Qy 5747 ATGACAGCAGAGCTGAGATTAATAAACCATTGTAAGCAATCCAGATGTCCTCAAT 5806
 Db 6157 ATTACACGAGAGTGTCCAAACTTAAGCAATTTGAAACATCAAGATGTTCTATGTC 6216
 Qy 5807 ACCATGATTTTTCCTCCCTTTTGTCTATGCTGATGCTGAGTGAAGAGTCTCTCTG 5866
 Db 6217 ATTGGGAATTTTTC-----TTTCTAACCCAGTGAAGGTTAGAAAGATTTATCT 6269
 Qy 5867 TGTGAGA-----TTAAGCCTGCTCTCTATGATGATGAGCAATGAT 5909
 Db 6270 GGTAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6329
 Qy 5910 GTGCTAAGGCGCATGAGA-----TGTTCTCATGAGCAAGATGCTGT 5953
 Db 6330 GTGCTAATCTGAGGCAAGTATGATGATGATGATGATGATGATGATGATGATG 6389
 Qy 5954 TGTACGTTTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATG 6013
 Db 6390 TGTAGG--TTTTGATTTACTCTTATATGCTGAGCTGATTCACACATGATGATG 6447
 Qy 6014 AGTCTGTTTCTTATAGATGATGATGATGATGATGATGATGATGATGATGATG 6073
 Db 6448 AGTCAAGTCTTATAGATGATGATGATGATGATGATGATGATGATGATGATG 6507
 Qy 6074 TGCCCTCTTCTTATAGATGATGATGATGATGATGATGATGATGATGATGATG 6133
 Db 6508 TGCCCTCTTCTTATAGATGATGATGATGATGATGATGATGATGATGATGATG 6567
 Qy 6134 CTGATTTAAATTA 6147
 Db 6568 CTGATTTAAATTA 6581

RESULT 5
 ABL61938
 ID ABL61938 standard; DNA; 6608 BP.

XX ABL61938;

XX 15-MAY-2002 (first entry)

XX Colon adenocarcinoma related gene sequence SEQ ID NO:275.

XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 XX stomach; lung; prostate; pancreas; carcinoma; antitumor; cancerous;
 KM cytoskeletal; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
 XX gene; ds.

OS Homo sapiens.

XX WO200194629-A2.

XX 13-DEC-2001.
 PD 30-MAY-2001; 2001MO-US10838.
 XX 05-JUN-2000; 2000US-209473P.
 XX 05-JUN-2000; 2000US-209531P.
 XX 18-SEP-2000; 2000US-23133P.
 XX 18-SEP-2000; 2000US-233617P.
 XX 20-SEP-2000; 2000US-234009P.
 XX 20-SEP-2000; 2000US-234034P.
 XX 20-SEP-2000; 2000US-234052P.
 XX 22-SEP-2000; 2000US-234509P.
 XX 22-SEP-2000; 2000US-234567P.
 XX 25-SEP-2000; 2000US-234923P.
 XX 25-SEP-2000; 2000US-234924P.
 XX 25-SEP-2000; 2000US-235077P.
 XX 25-SEP-2000; 2000US-235082P.
 XX 25-SEP-2000; 2000US-235134P.
 XX 25-SEP-2000; 2000US-235240P.
 XX 26-SEP-2000; 2000US-235637P.
 XX 26-SEP-2000; 2000US-235638P.
 XX 27-SEP-2000; 2000US-235711P.
 XX 27-SEP-2000; 2000US-235720P.
 XX 27-SEP-2000; 2000US-235840P.
 XX 27-SEP-2000; 2000US-235863P.
 XX 28-SEP-2000; 2000US-236028P.
 XX 28-SEP-2000; 2000US-236032P.
 XX 28-SEP-2000; 2000US-236033P.
 XX 28-SEP-2000; 2000US-236034P.
 XX 28-SEP-2000; 2000US-236109P.
 XX 28-SEP-2000; 2000US-236111P.
 XX 28-SEP-2000; 2000US-236842P.
 XX 29-SEP-2000; 2000US-236891P.
 XX 02-OCT-2000; 2000US-237172P.
 XX 02-OCT-2000; 2000US-237173P.
 XX 02-OCT-2000; 2000US-237278P.
 XX 02-OCT-2000; 2000US-237294P.
 XX 02-OCT-2000; 2000US-237295P.
 XX 03-OCT-2000; 2000US-237316P.
 XX 03-OCT-2000; 2000US-237425P.
 XX 03-OCT-2000; 2000US-237598P.
 XX 03-OCT-2000; 2000US-237604P.
 XX 03-OCT-2000; 2000US-237606P.
 XX 03-OCT-2000; 2000US-237608P.
 XX 01-NOV-2000; 2000US-244867P.
 XX 01-NOV-2000; 2000US-245084P.
 XX (AVAL-) AVALON PHARM.
 XX Young PE, Augustus M, Carter KC, Edner R, Endress G, Horrigan S;
 P1 Soppet DR, Weaver Z;
 XX WPI; 2002-188264/24.
 XX Screening for anti-neoplastic agent involves exposing cells to a
 PT chemical agent to be tested for anti-neoplastic activity, and
 PT determining a change in expression of a gene of a signature gene set -
 XX
 XX Claim 1; SEQ ID 275; 44pp; English.
 XX The present invention describes a method (M1) for screening for an
 CC anti-neoplastic agent. The method involves exposing cells to a chemical
 CC agent to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (i)
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening
 CC an anti-neoplastic agent, and can be used for producing a product which
 CC is the data collected with respect to the anti-neoplastic agent as a
 CC result of M1, and the data is sufficient to convey the chemical
 CC structure and/or properties of the agent. M1 can be used in the

CC test of cancer such as colon, breast, stomach, lung, thyroid,
 CC esophageal, ovarian, kidney, prostate or pancreatic cancer,
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilms' tumour.
 XX
 SQ Sequence 6608 BP; 2091 A; 1418 C; 1827 G; 1272 T; 0 other;

Query Match 34.8%; Score 2144.4; DB 24; Length 6608;
 Best Local Similarity 64.3%; Pred. No. 0;
 Matches 4215; Conservative 0; Mismatches 1796; Indels 543; Gaps 45;

QY 3 GCGCGGCGGAGTAGAGGCCACTGAGCCATGGCCGAGCGAGTCCACCGAGCGCGG 62
 DB 162 GGTAGGCGGGAGAGTGGGAGGAGCCATGGCGCCGGAGCTCCACCGAGCAGCG 221
 QY 63 AGCCCGAGCAGCGG-----GGGGGAGCGACACGCCGAGCGAGTGGTCTCAGTGGCCAT 119
 DB 222 AGCCCGAGCAGCGCGCGGAGGGAGCTCCACCGCGGCTGAGCCCGAGCGCGGCGC 281
 QY 120 GGGCCCGCAGCTGAGCCTCGGA-----GCAGCTGGAGACCCCGCGAGCGCGAC 170
 DB 282 GGGCCCTCGCGGCGGCGCGCGAGACACCCACCGCGACCCCGCATCGCTCGCTCGGAC 341
 QY 171 CCGCCACCAAGCTCCACAGAGAGATGGCCAGCTGTCTCTCAACGGCGTAGCTGAA 230
 DB 342 CCGCCACCAAGCTCTACAGAGAGATGGTCACTGCTCCACCATCAATGGCGTAGCTGAG 401
 QY 231 CAAGGAGATGTCCATGTCCCAAGAGGAAACACAGGAGG----- 268
 DB 402 CAAGATGAGCTCAGCTCCAGGAGGGTGACCTAAATGGCCAGAAAGGAGCGCTGAACGGT 461
 QY 269 -----GCAGGAGGAAGAGTCTGATGAGGATGTGGACACGCA 308
 DB 462 CAAGGAGCCCTAAACAGCCAGGAGGAAGAGTCAATGTCAAGGAGGTGGACAGAGA 521
 QY 309 GAGTCAGAGATGTGAGAGAAAGACCGAGTTGAAGAAATGGCGGCAACTCCACAGCT 368
 DB 522 GACTCTGAAGATGTGAGCGAAAGAGACTCCCGATAAAGAGATGGCTACTAAAGTCAGCGGTT 581
 QY 369 GTTGAAGATATCAAAAGGATGGCAGGAGAGACATCAGAAATTAATGAAACAGATCCCT 428
 DB 582 GTTCAAGACATCAGATGATGGCAGGAGGAG-----ACCGAATATCGAACAGATTCTCT 638
 QY 429 GCTTCAGAAAAAATGTGAAGAAATGGTACAGCTCTGAGTCCAGGCTTAATGATTT 488
 DB 639 TCTTCAGAAAGCAATTTAGAAGAGCTAACACACCCACTGAGTCCCAAGGCTAATGATTT 698
 QY 489 GCGTTCAGAAAGATTTAAATTTGTTGGTTTAAATTCAGGTTGAAGAGGATAAAT 548
 DB 699 GGAATTTAAGAAAGTGTAAAGTTGTTGGCTTTAAATTTCACTGTGAAAAAGGATAAGACA 758
 QY 549 GAAAGTTCAGATCTGTCCAACTACTCACTGTCAAGAGGATGAAGCGGAGGCGAGAA 608
 DB 759 GAGAGCCTGACACTGTCCAGTACTCACTGTGAGAAAGATGAGGGGAGGAGCAGCA 818
 QY 609 GCGTCTGTGCGAGTGTGAGACCAACCGAGGCGCCAGTGTGAGAGTGTGCGTGGAGAGTCA 668
 DB 819 -----GGGGCTGGCGGACCCAGGACCCCGAGCTTGGGCT-----GGAGAGCA 863
 QY 669 GCATCCAAAGAAAGTGTGAGCAATCCACAGAGAGCAAGGACACCTGTGAAGCAA 728
 DB 864 GCATCCAAAGAAAGCGCAACCCCAACAAATCTACAGAGAAACCCGAGAGAGACCTGGAAGCGT 923
 QY 729 GAAACAGAGCAGCAGAAATCCCGCTTCAAGCCGAATCTGATCAAGCGGCTGAGGAGAA 788
 DB 924 GAGCAAGCCACGAGAAATTTCTCCCGCAGCCGAATCTGGCCAGAGCT-----GGAGGAA 980
 QY 789 GCCAAAGATGAAGAGAGAAAAACAGAGAAAGAGCCCAAGTCCCGCAGAAATCCCCCG 848
 DB 981 TGCAAGAGAGAGAGAGAGAAACAGAAAGAGAACTAGCAAGTCTGAGATCTCCG 1040
 QY 849 AGCAGCCAGTCAACAGTAGAGACAAATCTCTCTTCAAGAAAGTTCTTCACTCAGGTTGG 908

DB 1041 ACTAGTCCCGTGCCAGTGAACAGGATCAACCTTCAAAAAATTTCTCACTCAAGTTGG 1100
 QY 909 GCGCGTGGCGCAAGAGACCAAGCTTCAAGAAATCAAAAGAGGATGATCTGGAATCTGCC 968
 DB 1101 GCGCGTGGCGCAAAAGAGACCAAGTTCAGGAAGCCGAGGAGGATGAGTGGAGTTCA 1160
 QY 969 GAGAGAGAAAGAGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1028
 DB 1161 GAGAGAGAAAGAGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1220
 QY 1029 CCAGCCTCGGAGG-----GCAGGAGCGCGAGAGAC 1061
 DB 1221 GTTGCCTTCGAGAACTGACCGCTCCGAGCAAGCCACCCACAGGAGCGCGGAGAAAGT 1280
 QY 1062 ACAGACCGAGCCAGGTTGTTCAGCAGACTACGAGAAAGTGGAGCTGCTTTTGGAGAGCCAG 1121
 DB 1281 GCCACGAGCGCGGTTATCAGCTGAATATGAGAAAGTTGAGCTGCCCTCAGAGGAGCAA 1340
 QY 1122 GTTGTGACCTGAGGAGTGTTCAGAGGAGAGTGTGCTCTTTTGGCAACGGAAGTGTCT 1181
 DB 1341 GTGAGTGGCTCGCAGGAGCCTTCTGAAGAGAAACCTGCTCGTTGGCGACAGAAAGTGTCT 1400
 QY 1182 GATGAGAGATGGAAGCCACCAAGAA--GTTGTTTCAGAGGCTCCACGTCAGCACCGTG 1238
 DB 1401 GATGAGAAATAGAAATGTCACCAAGAGAGGTTGTGGCCGAAAGTCCACGTCAGCACCGTG 1460
 QY 1239 GAGAGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1298
 DB 1461 GAGGAGAGAACGAGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1496
 QY 1299 GGAACAGGAGAAATCTTTGCCCCCTGAGAAATCTGGCTGAGCCCCCAGGAGGTCCTCCAGGAA 1358
 DB 1497 GAAACAGCAGGCTCTGTCAGCTGAGAAATGTTGTTGGAATGATGAGAGGAGGAGGAGGAG 1556
 QY 1359 GCTGAGCTGCTGAGGAGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1418
 DB 1557 GCCGAACTGCGCAAGGAGTGGTGAAGCTCAAGAAACGTTGTTTCCGAGAGGAGCCCT 1616
 QY 1419 ACTCAACTGACAGACCTTAAGTCTGAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1478
 DB 1617 ACACAGGAGTGAACCTGAGTCTGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1676
 QY 1479 GTGAGTGGAGTGGAGTGTCTCTCAGGAAAGAAATCAAGGTACAGGAGGAGGAGGAGGAGGAG 1538
 DB 1677 GTGAGTGGAGTGGAAATGCTGTCATCACAGGAGAGAAATGAAGCTGAGGAGGAGGAGGAGGAG 1736
 QY 1539 AAGAACTCTTCTGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1598
 DB 1737 AAGAACTCTTCTGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1796
 QY 1599 CGAGGAGGTTGGGAGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1658
 DB 1797 AGAG 1853
 QY 1659 GAGAGTGTCTGATGAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1718
 DB 1854 GACAGCCAGGAG 1913
 QY 1719 ACCACGTTCTGAGAGAAAGGCGCTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1778
 DB 1914 ATCACGTTCTGAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1973
 QY 1779 ACTACTTCGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1838
 DB 1974 GCTACTTCGAGAG 2033
 QY 1839 ATGGTGACACCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1898
 DB 2034 ATGGTGACCCAGAGAG 2093
 QY 1899 GAGAGGTCAGAG 1958

Db 2094 GACAAAGTCAAGACGCTACTTGTCTTCCACCGAGACACAGCTCTGAATGCAAGAA 2153
QY 1959 GAAGTCAAACTGTTGGTAGAGAACTAAAGCCAGAGAACCAAGCCGTAGGGTGGATCT 2018
Db 2154 GAAATGAAAGGAGACGTGGAGAGCGCAAGCCGGAAGAACCAAGCGCAAGTGGATCC 2213
QY 2019 TCAGTGTCTTGGGAGACATGATTTGTGTCCGATCATCAAGAGAGAGCAAGAGCA 2078
Db 2214 TCAGTATCTTGGGAGCTTATTTGTGTGGATCATCAAGAGAGAGCAAGAGAG 2273
QY 2079 TCCCTTTCAGATGATGAGAGGCGCAAGACACTGGAGGGGACAGTCAAGAGCAGAG 2138
Db 2274 TCCCTTTCAGATGAGAGGCGCAAGAGCAATGGAGAGACACCAAGAGCTGAT 2333
QY 2139 GAGCCAGACAAAGCAAGAGAGCCGGAACAGACGCTGTTCTGCGACGACCCAGAGAG 2198
Db 2334 GAGCCGAGAGAGAGCAAGAGAGCGGAGACAGCGGATCTTGTGTTCCCAAGACAT 2393
QY 2199 GACCAAGCGCAAGAGATTTCTCACCCGAGCCAGCGGAGAGCCCTTCCGAGGGAGGT 2258
Db 2394 GATCCAGGGCAGGAGATTTCTCCCGGAGCAAGCTGAGAGCCCTTACCGAAGGGAGG 2453
QY 2259 GTCTCACTTGGGAGTCAATTTAAAGATTAGTCACTCCAGAGAGAGAGAGAGAGAG 2318
Db 2454 GTTTCACCTGGAGTCAATTTAAAGATTAGTCACTCCAGAGAGAGAGAGAGAGAG 2513
QY 2319 CTGGAAGAGAGAGCCGAGAGC-----TCTAGTGAAGCAGTGTCTCACTGAG 2366
Db 2514 CTGGAAG 2573
QY 2367 ATCGAAGCAGATGAGAGAGATCTTGGGTTTCCATTAAAGATTCATCCCGAGAGCG 2426
Db 2574 ACTGAACCCGATGAAGAGATCTGGGCTCTCATCAAGAGATTTATCTTGAAGAGAG 2633
QY 2427 AAG 2486
Db 2634 AAG 2693
QY 2487 ATAAATGAG 2546
Db 2694 GCGAAG 2753
QY 2547 GAGAGAGAGAGAGAG-----AAGCCAGGGAGATACAGAGAGAGAGAGAGAG 2597
Db 2754 GAAAGGAG 2813
QY 2598 GGGGCTGTAGCTGTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2657
Db 2814 GAG 2873
QY 2658 GTCAATTGATGGAG 2717
Db 2874 GTCCGAG 2933
QY 2718 GCTTCCGTAACAG 2777
Db 2934 GCTTCAAG 2993
QY 2778 GTCACTGAAG 2834
Db 2994 GTATTGAG 3053
QY 2835 GAGGATGAAGATGCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2894
Db 3054 GAG 3113
QY 2895 GTGACAG 2954
Db 3114 GTGACAG 3173
QY 2955 GCCGAG 3014
Db 3174 GCTGAAG 3233

QY 3015 ACAAG 3074
Db 3234 ACAG 3293
QY 3075 GAGGAG 3134
Db 3294 GAG 3353
QY 3135 GTGCTGAG 3191
Db 3354 CTGCTGAG 3413
QY 3192 GAGGTAG 3251
Db 3414 AGACAG 3473
QY 3252 GAGCCGAG 3311
Db 3474 GTAGATGCTAG 3533
QY 3312 ACTCAG 3359
Db 3534 ACCCAG 3593
QY 3360 GTCCGAG 3401
Db 3594 GTAACTGATGAG 3563
QY 3402 CAGGCGGAG 3461
Db 3564 CAGGCTATCCCGCTGAG 3713
QY 3462 CCTTGAAG 3521
Db 3714 CCGGAG 3773
QY 3522 GACAGTAAG 3581
Db 3774 GAG 3833
QY 3582 GCTCAG 3641
Db 3834 GCAAG 3893
QY 3642 GAG 3701
Db 3894 CAATCAAG 3953
QY 3702 CCGGCTTGGAG 3752
Db 3954 TCCATTTGTCAAG 4013
QY 3753 GATGAG 3806
Db 4014 AAG 4073
QY 3807 ACTCAAG 3866
Db 4074 CCGGAG 4133
QY 3867 GAAAG 3911
Db 4134 AAGATGATGCTTGAAG 4193
QY 3912 GAGATGAG 3971
Db 4194 GAGATGATGAG 4253
QY 3972 GAAG-----TGACAG 4028
Db 4254 GAG 4313

QY 4029 CTGACACTTGACATGCCAGCTCAGAGGGGGAGGCACTGGGAAGCCTTGGAGGAGC 4088
DB 4314 CAGACAGTGAATGTGCCCATATAGATGGGGCAAGGAAGTCAGCAGCTTTGGAGGAGC 4373
QY 4089 CCTTCTC---TCCAGACCAAGACAAAGCAGGTTGTCATAGAGGTTCAAGTTCAAGCCTG 4145
DB 4374 CTTCTCTCCCTAGCTCAGAGGAGGAGTATGCAACCAAAATTCAGTTTCAGAGCTCT 4433
QY 4146 GACACAACTGCTCACTCAACAGCAGAAAGCTGTGGAAAA-----GGTCATAGAAACGGTT 4199
DB 4434 GAGGCATCATCTCTTAACAGCGGCTGCAGAGGAGGAAAAAGGCTTTAGGAGAACTGCC 4493
QY 4200 GTGATTTTCAGAGCAGGTGAAGTCCAGAGTGTCTAGGTGCACACTTTATTACAGCTGAG 4259
DB 4494 AACATTTTGAAGACAGGTGAACCTTTGGAGCCTGCAGCTGCACATTTTAGTTCTGGAAAG 4553
QY 4260 AAGTCTCTGCAACGGGTGGCCACTGTGACTCTTCAGCATGCAGAGGACAGGTACCCCTG 4319
DB 4554 AAATCTCTGAAAAAATGAAGACTTTTGGCGCTCATCCAGGGGAGATGCTGTGCCACA 4613
QY 4320 GGGCCTGAGTCTCAGGCAGAAATCCATCCCAATCATAGTAATCTCTGCTCTCTGAAAGCAC 4379
DB 4614 GGGCCGACTGTGAGGCAAAATCGACACCACTGATATCTGTCTACTACCAAGAAAGGC 4673
QY 4380 CTACATCTGACCTACAGCAGAAATGAAGCGCATCCAGAGAGCGCATCAGAGGAAGAG 4439
DB 4674 TTAAGTTCCGACCTGGAGAGAGAAAAACCATCACTGAAGTGGAAAGTCAGATGAATC 4733
QY 4440 GACAAGCCAGATGCTGCTGCTGATGCTGACGGCAAGCAGAGTACAGCAATCGAAAAAGTC 4499
DB 4734 GATGAGCAGTTGCTGCC-----AGGAGTCAAGTGAGTGTAGCAATG---AGGAT 4784
QY 4500 CTCAAGCTCAACCTGAGATCTCTGGAACTTTGAGAGTGAAGAGCAACAAGATTTGTCTGAAC 4559
DB 4785 TTAGAGCTCAAAATGGATTTTGGAACTTTGAGACCAAAAGCAGTAAACTTTGTCCAAAAC 4844
QY 4560 GTCAATTCAGACAGCCGTTGACCACTGCGACGTAC-----AGAAACAGCCCCCGAAACTCAT 4616
DB 4845 ATCATCCAGACAGCCGTTGACCACTTTGTACGTACAGAGAAACACAGCCACCGAAATGTTG 4904
QY 4617 GCTTATGATTACAGACCCAGGTTTC-----CTGCATGCAGGCTTGACAG 4660
DB 4905 ACCTCTGAGTTACAGACAGCTCAGCTGATAAAGCTCAGCCAGGAGCGCTGGACAG 4964
QY 4661 CAGGGAGCCCAACAGATGCTTGGCAAAATGAAGATGCCAAGATGAACACCCAGTGCC 4720
DB 4965 GAAACGAGAAAGAGGAGGAGAACTCAGGCTCTGCACAGGATGAACACCAATTAAT 5024
QY 4721 GCAGCCCA-----GAGAGGACTTGCAAGTCTGACCGTTCTGGA 4759
DB 5025 TCAGCCAAAGAGGAGTCAAGAGTCAACCGCAGTGGGCAAGCAGCATCTTGATATTTCCAAA 5084
QY 4760 GGCATGGCTCAGCCTCGGAAATGCTTGCCGCGCTTGCACTTGAAAGCGCGGTGTCAA- 4818
DB 5085 GACATGAGTAGAGCTCAGAAAAAGACCATGACTGTTGAGGTGAAGGTTCCACTGTAAAT 5144
QY 4819 ----- 4818
DB 5145 GATCAGCAGCTGGNAGAGTGTCTCTCCCTCCTGAGGAAGAGGAGGTGGAGCTGGNACA 5204
QY 4819 -----AGTAAGCATTTGAGAAAGCTGCCT 4840
DB 5205 AAGTCTGTGTCAGAAAGATGATGCTCATGCTTGTGTAGCAGAAAGAAATAGAGAACTCACTA 5264
QY 4841 CCTCAACCCCAAGA--TCCAAAGGAGCATGCTGCTGATGGCCCTCAGCTCCAAAGCTTA 4898
DB 5265 GTTGAACCGAAAGAAATGAAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 5324
QY 4899 GCCCAGCAGAGGCGAGTCTCTGGAACCTAAACCAAGAAATCCCCAGACACCAACCGGA 4958
DB 5325 GCCCTGGCTGATCTGATGCTCAGGAGGCTTAACCAAGAGTCCCCAGATACAAATGGA 5384
QY 4959 CCAAAGCTAACCGAGGAGGCGATCCCCCAAAAGTTGAGGTCCAGGAAGAAGAAATGTCT 5018

DB 5385 CCAAAACAAAAAGAGAGGAGGATGCCAGGAAGTAGAATTGCAGGAAGAAAAAGTGCAC 5444
QY 5019 ACCAAGTCACTCAAGAGAACAAAGGCCAGGCAAGAGGACCTGCAGGAGCCAAAGGGA 5078
DB 5445 AGTGAATCAGATAAAGGATCACACCCCAAGACACAGAGGAGTTACAGAAACAAGAGAGA 5504
QY 5079 GACCTGGCAGAACTCCTAAGATGTTAGTTG-----CTCAITGTATC 5117
DB 5505 GAATCTGCAAGTCAAGACTTACAGAACTTTAAACATCATGCAGTTAAACTCATTTGTCT 5564
QY 5118 ATCTGTAAGACCAAGATGTGAAACAAGTACAGAACAA-----GATGCTGCTTTGGGAC 5173
DB 5565 GTTGAAGAACCAAGATGTGAAGACAAGTAGTAGAAGAAATGAATGCTGCTGCTGAGA- 5623
QY 5174 CTTGAGACCAAGATTTTCAGAGCCCATAGATCCAGAGCAGGCGCTTCCAAATGATTTC 5233
DB 5624 CTGAAGCCAGTATTTTCAG-ACCTTGAGAAATTGGAGAGCAGGCACATCACTGATCTCA 5682
QY 5234 ACCAGTAGACACCCCGCAATTTCTGAGCTTTCTGCGGAGCTAGAGCAGCTACAATT 5293
DB 5683 TTTCTAGAGAGC-CCCTGACAATCTGAGGCTTTCATCAGAGCTAGAGCCATTTAACTT 5741
QY 5294 TCCCTGTTTCAAGACTGCCTTTGATTTGGCCCTTGTGATGCGCTCGGTGATTTCTAACATT 5353
DB 5742 TCCTCTTTTCAAGACCAACCTACAATTTTCCCTTGATAAC----- 5781
QY 5354 TCCTCGTTTCAAGACTGCCTTTGATTTGGCCCTTGTGATGCGCTCGGTGATTTTCGGATTTA 5413
DB 5782 -----CATATAAATCTCGATTTA 5799
QY 5414 AGTCTCGCTGCTTCAACCTGGAAAC-CAATTTGCCATACCTAGTTTCCACTTCTCAAACT 5472
DB 5800 AGTCTCTAAATCTTAACCTGGAACTGGAGTTGGCAATACCTAGTTCTGCTTCTGAACT 5859
QY 5473 GAGCATCTCTCTTATGATTTATATGATGATGATGATGATGATGATGATGATGATGATGAT 5532
DB 5860 GGAGTATCATCTTTACATATTTATATGATGATGATGATGATGATGATGATGATGATGATGAT 5916
QY 5533 TGTATATTTTTTCTAACGTTT-----AAGCACATGCTTTTTTGTATTA 5575
DB 5917 TGTATATTTTTTCTTAATGTTTAAAGAAATGTGAGGATATACATGCTTTTTTGTATCA 5976
QY 5576 TGAATATATAACGGGTGTGCAGCCATAGCAGCTTTGAAAGCTCCAGCCTCAACTG 5635
DB 5977 CACAGTATATGATGGGCATGTGCCATAGTGCAGGCTTGGGAGCTTTAAGCCTCAGTTA 6036
QY 5636 TAACTGTCAGCAACAGATAACATTC-----CTGGCAAGAGAGACAAGTCTTTTT 5686
DB 6037 TATAACCCCAAAAAACAGAGCCTCCTAGATGTAAACATTCCTGATCAAGGTACAATTTCT 6096
QY 5687 TAAAGTTTACTGATGCTTAGATCTGTGGGCTTCTAGTCTCTGTAAGAGTGGTTGTTTCTCT 5746
DB 6097 TAAAAATCTACTAATGATGAGGTCCATATTTAGTGTACTCTGAAATTTGGTCTCTTCTCT 6156
QY 5747 ATGCACAGCGAGCTCAGAAATAAAAAACCCCATTTTGAACATCCAGGATGTCCCAATATT 5806
DB 6157 ATTACAGGAGTGTGCCAAAACTTAAAGAGATTTTGAACATACAGAAATGTTCTATTGTC 6216
QY 5807 ACCATGATTTTTTCCCCCTTTTGTCTAATTCAGTCCAGTCCAGTTGGAAGAGTCTCCTCTG 5866
DB 6217 ATTGGAAATTTTGC-----TTTCTAACCCAGTGGAGGTTAGAAAGAGTATATTTCT 6269
QY 5867 TGTTCAGA-----TTAAGCCCTGTCTCTTAATGATATGCAGAAATGAGT 5909
DB 6270 GGTAGCAAAATTAACCTTTTCACTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 6329
QY 5910 GTCCCTAAGGCCATGAGA-----TGTTTCTTAATCAGAGGAATCTGT 5953
DB 6330 GTGCTTAACTCCTGAGCAAGTAGTGAATATGTTTTATATGTTATGAAGAAAGAAATGTT 6389
QY 5954 TGTACGTTTTTTGATGTTACTCTTCTATGCTGGACCGAAATTCATATGCAGATCGAAGTG 6013

Db 6390 TGTAA--TTTGTGATTCTACTCTTATATGCGAGCTGCATTCACACATGGCATGAATA 6447

QY 6014 AGTCCTGTTCTTTACAGATGGTATTTGATAGTACTGAGTTGTCTGTGTATATCTG 6073

Db 6448 AGTCAGGTTCTTTCAAAATGGTATTTGATAGTACTGAGTTGTGTGTGTGATATTTG 6507

QY 6074 TGGCCCTCTTTAGAAACAATGGCATTTATGTCCTTTGATAAATGTGATTTGACAA 6133

Db 6508 TGGCATTCCTTTAGAACAAATGTGCAACACATTCATTTGGATAGTTGTATTTGACGA 6567

QY 6134 CTGATTTAAATPAA 6147

Db 6568 CTGATTTAAATPAA 6581

RESULT 6
AAV23545
ID AAV23545 standard; cDNA; 6605 BP.

AC AAV23545;

DT 13-JUL-1998 (first entry)

XX Human gravin polypeptide coding sequence.

XX

XX Gravin, kinase anchoring protein; type II regulatory subunit; PKA; PKC;
KW CAMP-dependent protein kinase; protein kinase C; autoimmune disease;
KW Myasthenia gravis; nicotinic acetylcholine receptor; ss.

XX Homo sapiens.

XX OS

XX Key Location/Qualifiers
FT 192..5534
FT CDS /*tag= a

XX US5741890-A.

XX 21-APR-1998.

XX 19-DEC-1996; 96US-0769309.

XX 19-DEC-1996; 96US-0769309.

XX (UYOR-) UNIV OREGON HEALTH SCI.

XX Klaubek TM, Nauert JB, Scott JD;

XX WPI, 1998-260552/23.

XX P-PSDB; AAW53863.

XX New polypeptide fragments of protein kinase binding protein -
PT are useful for the study of modulation of action between gravin and
PT protein kinase(s)

XX Example 1; Column 19-32; 32bp; English.

XX This sequence encodes the human gravin polypeptide, which represents
XX a polypeptide of the invention. The polypeptides are fragments capable of
XX binding to type II regulatory subunit of CAMP-dependent protein kinase
XX (PKA). Gravin is a kinase anchoring protein that binds to type II
XX regulatory subunits of PKA and protein kinase C (PKC). Gravin is also an
XX antigen of the autoimmune disease Myasthenia gravis (MG), where a patient
XX develops antibodies against their own nicotinic acetylcholine receptors.
XX The polypeptides are useful for providing analogues of gravin in the
XX study of the modulation (e.g. blocking, inhibiting and stimulating) of
XX interactions between gravin and kinase. The peptides are involved in the
XX modulation of gravin-kinase interactions.

XX Sequence 6605 BP; 2091 A; 1416 C; 1826 G; 1272 T; 0 other;

Query Match 34.4%; Score 2118.8; DB 19; Length 6605;
Best Local Similarity 64.2%; Pred. No. 0;
Matches 4206; Conservative 0; Mismatches 1802; Indels 546; Gaps 46;

QY 3 GGCCGCGCGGAGTGAAGAGCACTGAGCCATGGGCGCAGAGCTTCCACCGAGCAGCG 62

Db 162 GGTAAAGCGCGGAGAAAGTGGAGAGCCATGGGCGCCGGAGCTCCACCGAGCAGCGC 221

QY 63 AGCCCGAGCAGCGCG--GCGGGAGGAGACCGCGAGCGAGCTGTGTCTAGTGGCAT 119

Db 222 AGCCCGAGCAGCGCGCGCGAGGGAGCTCCACCGCGCTAGCGCCGAGCGCGCGC 281

QY 120 GGGCCCGCAGTGAAGCTTGGGA-----GCACCTGAGAGCCCCCGCGAGCGGAC 170

Db 282 GGGCCCTCGCGCGAGGCGGCGCGAGACCAACCGCGAGCCCCCGCATCGCTCGAC 341

QY 171 CCGCGCAGCAAGCTCCACAGAAAGATGGCAGGTCTTCTGTCAAGCGGTAGTGA 230

Db 342 CCGCGCAGCAAGCTCTCAGAGAAATGTGTAGCTGTCCACATCAATGGGTAGTGA 401

QY 231 CAAGAGATGTCCATGTCCAGAGGAAACCAAGAGG----- 268

Db 402 CAAGATGAGCTCAGCCTCCAGAGGAGTGACTTAATGCGCAAGAGAGCGCTGAACGCT 461

QY 269 -----GCAGAGAAAGATCGTTGATGAGATGTTGACAGCA 308

Db 462 CAAGAGCGCTTAAACAGCCAGAGAGAGAAAGATCAATGTCAAGAGGTTGACAGAGA 521

QY 309 GAGTCAGAAAGTGAAGAGAAAGAACCGAGTGAAGAAATGGCGCACTCCACAGCT 368

Db 522 GACTTGAAGATGTGAAGAGAAAGATCTCCGATTAAGATGGCTACTTACAGCGTT 581

QY 369 GTTGAAGATATCAAGAAAGATGGCGAGAGAGACATCAGAAATATTTGAACAGATCCCT 428

Db 582 GTTCACGACATCACAGATGATGGCAGAGAGAGA---ACGAAATATTCGAACAGATTCCT 638

QY 429 GCTTCAGAAACATGTGGAAGAAATGTACAGCTGTGAGTCCCGGCTAATGATGTT 488

Db 639 TCTTCAGAAACCAATTGAAGAGCTAACACACCACATGTAGCTCCAGGCTAATGATATT 698

QY 489 GGCCTCAAGAAAGATTAATTAATTTGTTGTTTAAATTCAGCGTGAAGAGATTAAT 548

Db 699 GATTTAAGAAAGGTGTTTAAGTTGTTGCTTTAATTTCACTGTGAAGAAAGATTAAGCA 758

QY 549 GAAAGTCAGATCTGTCCAACTACTCTGTCAAGAGATGAAGGAGGAGGCGCAGAA 608

Db 759 GAGAAAGCTGACACTGTCCAGTACTCTGTGAAGAAAGATGAAGGAGGAGCAGCA 818

QY 609 GCCTTGTGGAAGCTGGAGACCAACAGAGCGCCAGTGTGAGACTGCCGTGGAGAGTCA 668

Db 819 -----GGGCGTGGCGACCAACAGAGCCCAAGCTTTGGGGCT-----GAGAGACCA 863

QY 669 GCATCCAAAGAAAGTGAAGCTGAGCAATCCACAGAGAAAGCAAGAGCACCCTGAAGCA 728

Db 864 GCATCCAAAGAAAGGGAACCCAAACATCTACAGAGAAACCGAAGAGACCTTGAAGCT 923

QY 729 GAACAGAGCAGACAGAAATCCCTTCAGCGCAATCTGATCAAGCGCTGAGAGAA 788

Db 924 GAGCAAAAGCAGCGAATTTCTCCCGACGCAATCTGGCGCAAGCAGT---GAGAGAA 980

QY 789 GCCAAAGATGAAGAGAAAGAAAGAAAGAGAGCCCACTAATGCCAGATTCCTCCG 848

Db 981 TGCAAAGAGGAGGAGAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1040

QY 849 ACCAGCCAGTCAAGTGAAGCAATCTTCTCAAGAAAGTTCTTCACTCAGCGTTGG 908

Db 1041 ACTAGTCCGTACAGATGAAGAAAGAGATCAACCTTCAAAAATTTTCACTCAAGGTTGG 1100

QY 909 GCGGCTGCGCGCAAGAGCCAGCTTCAAGAAATCAAGAGAGATGTGAGAAATCGCC 968

Db 1101 GCGGCTGCGCGCAAAAGACAGTTTCAGAAAGCGGAGAGAGATGAAGTGAAGCTTCA 1160

QY 969 GAGAAAGAAAG 1028

Db 1161 GAGAAAGAAAG 1220

3354 CTGCTGGACCGGTCGAGCAAGATGCTTCAGCCTGTGACAGAGACGAGGACGAA 3413
QY 3192 GAGGTAGAGAGACTCCGAAGTGTGCTTGGAGAAAGAGAGACGTTATGCCGAA 3251
Db 3414 AGACAGAGAGAGAGGCTGAAAGCGTCGGGTCTGAAAGAAAGACGAGTAGTGTGAAA 3473
QY 3252 GGACCCGTGACGAGAGCTGAGAGCTTGTGACAGAGGGCTGTGACGTGACGAGCT 3311
Db 3474 GTAGATGCTCAGAGGACAAAGCTGAGCCTTTTACAGAGAGAGAGAGTGTGGGACACC 3533
QY 3312 ACTCAGAGAGACCTTGG---AAGTCTCTGAAGTCAAGGAGATGAGAC-----CAT 3359
Db 3534 ACCCAGAGAGAGCTTGAAGAGCTCTCAAGTCAAGAGAGATGAGTCCAGAGCTT 3593
QY 3360 GTCCGACGTGCCA-----GGTTATCAGCTCCAGCAGCTGATGGA 3401
Db 3594 GTAACCACTTGTCAAGCCGAAACCTTACGTGGGGTAAATCAAGAGAGTGTGTGGA 3653
QY 3402 CAGGCGGTGCCCCCTGAGTCAATCCGAAACCTTGACAGACAGAGAGCAAAATGGAACACT 3461
Db 3654 CAGGTATATCCCTGAGTCTGGTGAACCCCTACAGACAGTGAATGATGGAAGCACCC 3713
QY 3462 CCTTAGCAGATTCAGACCTGACATGGGACACAGCAAGATGAAACATTTGACAGCCAG 3521
Db 3714 CCGGTATCCGACTTTTGAAGCACACAGCACACCCAGAAAGAGAGATTTGTGAAATCCAT 3773
QY 3522 GACAGTAAAGCCACTGACCTGTGACAGGACATCAGGTCAAGAAAGAGAGGCGCTACT 3581
Db 3774 GAGGAGAAATGA---GATGATCTGGTACAGTCAAGGGGACAGAAAGCAGAGCACTTCT 3830
QY 3582 GCTCAGAAAGAGAGACCTTGCACACTACTAATATTTCCAGCCCAAGAAAGATGGG 3641
Db 3831 GCACGAAAGAGAGAGGCTTCCAGACCTTCCAGTTTGTGTTCCAGGAAAGAACTTAAAGAA 3890
QY 3642 GAAAGACCGAGAGAGATTTCTTGAACCTACAGACAGAAAGACTTACTGCTCAGCCGTG 3701
Db 3891 CAATCAAGATGGAAGACACTCTTAGAGCATACAGATTAAGAGGTGTGATGGAACCTGTA 3950
QY 3702 CCGGTTCTGGCAAGACTGAGGTGGGTCAAGAGGGTGA-----GGTTGACTGGTTG 3752
Db 3951 TCCATTCTGTCAAGACTGAGGGGACTCAAGAGGCTGACCAATGCTGATGAGAAACCC 4010
QY 3753 GATGAGAGAAAGACTCAAGAGAAAGACAGAGGTGTTGTACTCTGG-----ACCCAC 3806
Db 4011 AAAGCGTACCAATTTTTCAGAGACTTGGGGGTCTATAGACACAGGCAATTAACGTCACT 4070
QY 3807 AGTCAAAAGGCTGCTGATGTGACATATGACATGAAAGTATGGAGTGGCCGGGTGTGACG 3866
Db 4071 CGGGAAGAGTCACTGAGAGTGGCTTTAAAGGTGAAGGAGACAGAAAGACTGAATGTAAA 4130
QY 3867 GAAAAAGAGATGATGAACTGACAG-----TCTTAAGCTTGAAGAGGA 3911
Db 4131 AAGATGATGCTCTGAACTGACAGAGTCAAGCTAAGTCTCTCAATCCCGGTGAGAGA 4190
QY 3912 GAGATGAGAACTGAGCTTGAAGAAAGAGAAAGGAGACAAACCCAGAGAAATGTAGTGA 3971
Db 4191 GAGATGATGATGTTCAAGTCAAGAGGAGAAACAGAGACAGACCAACCATATGATGAA 4250
QY 3972 GAAG---TGACAGAGAAACAGCCGCTCTGACATGAAGAACTTACCGGAAGCCAGTC 4028
Db 4251 GGAAGCTTGAAGCAAGAAAGCTGTTACCTGATCTGAAGAGGTCAGTAAAGAGCTCTTC 4310
QY 4029 CTGACACTTGACATGCCCCAGCTCAGAGAGGGGAAAGGCACTGGGAAAGCTTGAAGGAAC 4088
Db 4311 CAGACAGATGAATGTGCCCTCATATAGATGGGCAAGAGAGTCAAGCTTGAAGGAAGC 4370
QY 4089 CCTTCTC---TCCAGACCAAGACAAAGTGTGATGAGAGTTGAAGTCAAGCTTG 4145
Db 4371 CCTCTCCCTGCTAGGTGAAGGAGGAGCATATGACCAAAATTCAAATTCAGACCTCT 4430
QY 4146 GACCAACAGCTCAACACAGAGAGAGTGTGAAAA-----GGTCAATGAAAGCGTT 4199
Db 4431 GAGGATCATCTTAAACAGGCGCTGACAGAGAGAAAGGTCTTAAAGAGAACTGCC 4490

QY 4200 GTGATTCAGAGACAGGTGAAGATCCAGAGTGTAGTGACACTTATTAACAGCTGAG 4259
Db 4491 AACATTTTGAAGACAGGTGAAGCTTGGAGCTGACAGTGACACTTATGTTCTGGAAGAG 4550
QY 4260 AAGTCTCTGCAACGGGTGGCTGAGTCTTTCAGATGACAGAGAGACAGGTAACCCCTG 4319
Db 4551 AATCTCTGAAAAAATGAAGACTTTGGCCGCTCATCAGGGGAAAGATGTGTCCTCCACA 4610
QY 4320 GGGCTGAGTCTCAGGACGAATCCATCATATGTAATCTCTGCTCTGAAAGAC 4379
Db 4611 GGGCCCACTGTCCGAGAAATTCAGACAGATGATATCTGTACTACCAAGAAAGGC 4670
QY 4380 CTATATCTGACCTAACAGAGAAATGAAGCATCCAGAGAGAGCATCAGAGAGAG 4439
Db 4671 TTAAATTCGACCTGGAAGGAGAAACACATCACTGAATGGAAGTCAATGAAATGTC 4730
QY 4440 GACAAAGCAGATGCTGTCTGATGCTGACGAGCAAGAGATGACGAATGAAAAATGC 4499
Db 4731 GATAGCAGGTTCCTTGC-----AGGAGTCAAAAGTGAATGATGACATTTG---AGGAT 4781
QY 4500 CTCAAGCTGAACCTGAGATCTTGAACCTTGAGATGAAGCAACAGATTGTCTGAAC 4559
Db 4782 TTAGAGCTGAAATGGGATTTTGGAACTTGAGACCAAAAGCAGTAAACTGTCCAAAAC 4841
QY 4560 GTCAATTCAGACAGCCGTTGACACTTGCACAGTAC---AGAAAGCCCGGAAATCAT 4616
Db 4842 ATCATTCAGACAGCCGTTGACAGATTTGTATCAAGAAAGACCCAGAAATGTTG 4901
QY 4617 GCTTATGATTCACAGACCAGCTTC-----CTGCATGACAGCTTGAACAG 4660
Db 4902 ACGTGTAGTTACAGACACAGCTCAGCTGATTAAGCTGACACCGAGAGCTGTGACAG 4961
QY 4661 CAGGAGCCCAACAGATCTGGAACAAAATGAATGCCAAGTGAACACCCAGTGC 4720
Db 4962 GAAACGAGAAAGAGAGAGAGAACTCAGGCTCTGACAGAGATGAACACCAATTACT 5021
QY 4721 GCAGGCCA-----GAGAGACTTGCAGTCCGACCGTTTGA 4759
Db 5022 TCAGCAAGAGAGAGTCAAGTCAACCGCAGTGGACACAGCATTTGATATTTCCAA 5081
QY 4760 GGCATGGGCTCAGCTCGGAATGCTTGGCCGCTTGGACATTTGAAGGCGCGGTGCA- 4818
Db 5082 GACATGATGAAGCTCAGAAAGACATGACTGTGAGATGAAGATTCACACTGTAAAT 5141
QY 4819 ----- 4818
Db 5142 GATCAGAGCTGGAAGAGTGTCTCTCCATCTGAGAAAGAGAGGTGAGCTGGAACA 5201
QY 4819 -----AGTAACATTTGAAAGCTGCT 4840
Db 5202 AAGTGTGCAAGAGATGATGATGATGCTTTGAGCAAGAAATGAGAAAGTCACTA 5261
QY 4841 CCTCAACCCAAAGA--TCCAAAGAGAGATGCTGTGAGCCCTCAGACTCAAAAGCTTA 4898
Db 5262 GTTGAACCGAAAGAGATGAAGAAAGGTATGATGTGATGACCTGGAACACAGAACTCA 5321
QY 4899 GCCCAGGAGAGGCCAGTCCCTTGGAAAACCTAACCAAGAAATCCAGACCAACGGA 4958
Db 5322 GCCCTGGCTGATATGATGCTCAGAGAGGCTTAAACCAAGAGTCCCGCATACAAATGGA 5381
QY 4959 CCAAGCTTAACCGAGAGAGGAGATCCCCAAAGATTTGAGGTTCAGAGAAAGAAATGTCT 5018
Db 5382 CCAAAACCAAAAGAGAAAGAGAGATGCCAGAGAAATGAATTTGCAAGAAAGGAAAGTGCAC 5441
QY 5019 ACCAAGTCACTCAAGAGAGAACAGGCCCCAGGACAGAGAGGACCTGACAGAGCAAAAGGA 5078
Db 5442 AGTGAATCAGATTAAGCGATACACCCAGACAGAGAGGATTAACAGAAACAAAGAGA 5501
QY 5079 GACCTGGCAGATCTTAAGATGTTAGTTG-----CTCATTTGATC 5117
Db 5502 GAATCTGCAAGAGTCAAGAACTTAACGAATCTTAAACATCATGACAGTTAACTCATTTGCT 5561

```
QY 5118 ATCTGAAGACCAAGATGTGAACCAAGTACAGAACAA---GATGCTGCTGTGGGAC 5173
Db 5562 GTTTGGAGAGACCAAGATGTGAAGCAAGTAGTAGAGAAATGAATGCTGCTGTGAGA- 5620
QY 5174 CTTGAGACCAAGATTTTCAGAGCCCATGAGATCCAGAGAGAGGGCCGTCCAATGATTTCC 5233
Db 5621 CTGAAGACCAAGATTTTCAGA-ACITTTGAGAATTTGGAGAGAGGCACATCAACTGATCTCA 5679
QY 5234 ACCAGTAGAGCACCCCGACAAATCTGAGGCTTCATCGGAGCTAGAGCCAGCTAAACATT 5293
Db 5680 TTTCTAGAGAGC-CCCTGACAATCTGAGGCTTCATCAGGAGCTAGAGCCATTTAAACATT 5738
QY 5294 TCCTCGTTTCAAGAGCTGCTTTGATTTGCCCCCTTGATGCGCGTCCGTGATTTCTTAACATT 5353
Db 5739 TCCTCTTCCAGAGCAACCTACAAATTTCCCTTGATAAC----- 5778
QY 5354 TCCTCGTTTCAAGAGCTGCTTTGATTTGCCCCCTTGATGCGCGTCCGTGATTTCTGGAATTA 5413
Db 5779 -----CATATAAATCTGATTTA 5796
QY 5414 AGTCTGCGCTTCTCAACCTGGAAAC-CAATCTGCCATACCTAGTTCCACTTCTCAAACT 5472
Db 5797 AGTCTCAAAATCTTAACTCTGGAACTGAGTTGGCAATACCTAGTTCTGCTTCTGAACT 5856
QY 5473 GGAGCATCTCTCTTATGATTTATATATGATGTTTATGTTAGTCTCTCTCTGCTACCTAT 5532
Db 5857 GGAGTATCATCTTTACATATTTATATGATGTTTAAAGTAG---TCCTCTGATCTAT 5913
QY 5533 TGTATATTTTTTTCTAACTGTT-----AAGCACATGCTTTTTTGATTA 5575
Db 5914 TGTATATTTTTTTCTTAATGTTTAAAGAAATGTGCAGGATCTACATGCTTTTTTGATCA 5973
QY 5576 TGCATATATAAGGGTGTGAGCCATAGCGAGCTTTGAAAAGCTCCAAAGCCTCAACTG 5635
Db 5974 CACAGTATATCATGGGCGATGTGCCATAGTGCAGGCTTGGGAGCTTTAAGCCTCAGTTA 6033
QY 5636 TAACCTGCAGCAACAGATAACATTC-----CTGGCAAGNAGAGACAAGTCTTTTT 5686
Db 6034 TATAACCCACAAAACAGAGCCTCTAGATGTAACATTCCTGATCAAGGTACAATTCCT 6093
QY 5687 TAAAGTTTACTGATGCTTAGATCTGTGGGCTCTAGTCCCTGAAAGTGCTGTTTTCCT 5746
Db 6094 TAAATTTCACTAATGATGAGTCCATATTTAGTGTGACTCTGNAATTTGTCATTTCT 6153
QY 5747 ATGCACAGAGCTCAGAAATAAAACCCCATTTTGAACATCCAGGATGTCCCAATATT 5806
Db 6154 ATTACAGGAGTGTGCCAAAACCTAAAGAGCATTTGAACATACACAGAAATGTTCTATTGTC 6213
QY 5807 ACCATGATTTTTTCCCCCTTTTGTCTAATCCAGTCCAGGTTGGAAGAAGTCTCTCTG 5866
Db 6214 ATTGGGAAATTTTGC-----TTTCTAACCCAGTGGAGGTTAGAAGAAGTATATTCT 6266
QY 5867 TGTGAGA-----TTAAGCCCTGCTCTCTTAATGATATGACAAATGAGT 5909
Db 6267 GGTAGCAATTAACCTTTACATCCTTTTCTACTGTTATGTTGTTGGACCGATAGT 6326
QY 5910 GTGCTTAAGCCCATGAGA-----TGTTTCTTAATGCAGAGGAATCTGT 5953
Db 6327 GTGCTTAATCCTGAGGCAAGTAGTGAATATGTTTATATGTTATGAAGAAAGAAATGT 6386
QY 5954 TGTACGTTTTTTTGTATGTTACTCTTCTATGCTGACCGAATTCATATGCAGATCGAAGTG 6013
Db 6387 TGTAAAG-ATTTTGTATTTCTACTCTTATATGCTGGACTGCATTCACACATGCGCATGAATA 6444
QY 6014 AGTCTCTGTTCTTACAGATGGTATTTTGTATAGATCTGGAAGTTTGTCTGTTTATATCTG 6073
Db 6445 AGTCAGGTTCTTTACAAATGGTATTTTGTATAGATCTGGAATTTGTTTGGCCATATTG 6504
QY 6074 TGCCCTCTTCTTAAAGAACAAATGTTGCAATATGTTTCTTTTGGATAAAATTTGATTTGACAA 6133
Db 6505 TGCCATTTCTTTTAAAGAACAAATGTTGCAACACATTTCAATTTGGATAAGTTGTGATTTGACGA 6564
QY 6134 CTGATTTAAATAAA 6147
```

```
Db 6565 CTGATTTAAATAAA 6578
RESULT 7
AAA74903
ID AAA74903 standard; cDNA; 6605 BP.
XX
AC AAA74903;
XX
DT 26-JAN-2001 (first entry)
XX
DE Human gravin coding sequence.
XX
KW Human; gravin; PKA RII binding site; myasthenia gravis;
KW kinase anchoring protein; CAMP dependent protein kinase; ss.
XX
OS Homo sapiens.
XX
PH Key Location/Qualifiers
FT CDS 192..5534
FT /*tag= a
FT /product= "gravin"
XX
PN US6090929-A.
XX
PD 18-JUL-2000.
XX
PF 19-DEC-1997; 97US-0994570.
XX
PR 19-DEC-1996; 96US-0769309.
XX
PA (UYOR-) UNIV OREGON HEALTH SCI.
XX
PI Klauck TM, Scott JD, Nauert JB;
XX
DR WPI; 2000-523763/47.
XX
P-PSDB; AAB15380.
PT Novel polynucleotides useful for detecting gravin in patients suffering
PT from Myasthenia gravis encodes CAMP-dependent protein kinase-binding
PT polypeptide and protein kinase C-binding polypeptide of gravin -
XX
PS Example 1; Column 23-36; 34pp; English.
XX
CC The present sequence is the coding sequence for human gravin. Gravin is
CC an A-kinase anchoring protein (AKAP) which is involved in the
CC localisation of CAMP dependent protein kinase A (PKA) via interactions
CC between the RII binding region and the PKA regulatory subunit RII. Gravin
CC is also an antigen found in myasthenia gravis sufferers, and it is
CC thought that antibodies to it may be useful in modulating the binding of
CC PKA, and thus aid in the treatment of the disease. This sequence was
CC isolated by first screening a human umbilical vein endothelial cell cDNA
CC library with serum from a myasthenia gravis patient, and then searching a
CC human heart cDNA library for sequences resembling the isolated sequence.
CC This was done because the first sequence obtained was shown to be shorter
CC than the full length coding sequence.
XX
SQ Sequence 6605 BP; 2091 A; 1416 C; 1826 G; 1272 T; 0 other;
Query Match 34.4%; Score 2118.8; DB 21; Length 6605;
Best Local Similarity 64.2%; Pred. No. 0;
Matches 4206; Conservative 0; Mismatches 1802; Indels 546; Gaps 46;
QY 3 GCGCCGCGGAGTAGAGAGCCACTGAGCCATGGGCGCAGGAGTTCACCCGAGCAGCG 62
Db 162 GGTAGTCGCGGAGAGAGTGGAGAGGAGCCATGGGCGCGGAGTTCACCCGAGCAGCG 221
QY 63 AGCCCCGAGCAGCGCG---GCGGGGAGCAGACGCGCCGAGGAGTGGTCTCAGTGGCCAT 119
Db 222 AGCCCGGAGCAGCGCGCCCGAGGGAGCTCCAGCGGCTGAGCCCGAGCCAGCGCGGC 281
QY 120 GGGCCCGCAGCTGAAGCCTTCGGGA-----GCAGCTGGAGACCCCGCGCAGCGGAC 170
```


QY 2259 GTCTCCACTGGGAGTCATTTAAAGATTAGTCACTCCAGAGAAAAAATCCAAGTCAAAA 2318
DB 2454 GTTTCCCACTGGGAGTCATTTAAAGGTTAGTCACGCCAAGAAAAAATCAAAGTCCAAAG 2513
QY 2319 CTGGAAGAGAAAGCCGAAGAC-----TCTAGTGTAGAGCAGTGTGTCACCTGAG 2366
DB 2514 CTGGAAGAGAAAGCCGAAGACTCCATAGCTGGGTCTCGTGTAGAACATTCACATCCAGAC 2573
QY 2367 ATCGAAACCGAGTAGAGAAAGATCTTGGGTTTCCATTAAAGAAATTCATCCCGGACGGCGG 2426
DB 2574 ACTGAACCCCGGTAAAGAAAGATCTGGGTCTCAATCAAGAGTTTATTCTCGGACGAAGG 2633
QY 2427 AAGNAAAGGCGACAGCGGAGCAAGACACACCTGTGTGGAAGACTCAGGSCCAGTGGAG 2486
DB 2634 AAGNAAAGGCGAGATGGGAAACAAGAACCAAGCCCTGTGTGAAGACGCGAGGSCCAACAGGG 2693
QY 2487 ATAAATGAGGACGACCCCTAATGTCCAGCCGTCTGTCTGTCTGTGATATTAATGTCAGTG 2546
DB 2694 GCCAACGAGATGACTCTGATGTCCCGCCGTGTCTCTGTCTGTGATATGATGCTGTA 2753
QY 2547 GAGAGGAGAGATGG-----AAGCCGAGGGGAATACGGAGCTGCCCGAGCTGCTG 2597
DB 2754 GAAAGGAGAAAAATGGAGGCACAGCAAGCCCAAAAAGCGCGCAGAGCAGCCGACCAAG 2813
QY 2598 GGGCTGTGTACGTTCGAGGAGCTCAGTAAAGACTCTGTGTCCACACTGTGTAGTGTGCGCA 2857
DB 2814 GCAGCCACTGAGGTGTCCAGGAGCTCAGCGAGAGTCAGGTTCATATGATGCGCAGCAGCT 2873
QY 2658 GTCAATGATGGACCGAGGAGTCAACAGTGTGCAAGCGGTCTCTCTGTGATATCC 2717
DB 2874 GTCGCTCAGCGGACGAGGGGAGTCACTTATTTAAAGAAAGGTCTCTCTGTGATATCT 2933
QY 2718 GCTTCGTTAAGACCTCTTTGAACACACAGCGGGGAAGCCATGCCACCTGTTGGAAGAG 2777
DB 2934 GCTTCAGTGACAGAACCTCTTGAACAGTAGAAGCTGAGCCGCACTGTTAACTGAGAG 2993
QY 2778 GTCACTGAAAGACATCATTCG---AGAGAAACTCTGTGTCTACCCAGAGCTTACCA 2834
DB 2994 GTATTGGAAGAGAAAGTAATTGCAAGAGAAAGAACCCCGCGGTACTGAAACCTCTGCCA 3053
QY 2835 GAGGTTAAGATGCCATGACGACATGGTCAACAGTCAAGTGAAGTTCACCTCAGAGCT 2894
DB 3054 GAGAACAGAGGCCCCGGGGCGACACGGTGTGTAGTGAAGCGGAATGACCCCCGAGCT 3113
QY 2895 GTGACAGCCACAGAGAGCTCTCGTACTCAAGAAAGTTACCGAAGCATCGGG 2954
DB 3114 GTGACAGCTGCAGAACTGCGAGGCCATTGGGTTCCGAGNAGNAAACCGAAGCATCTGCT 3173
QY 2955 GCCGAGAGACCAAGACATGGTGTCCGCAATTTCCAGCTGACTGACTCCCCAGACACC 3014
DB 3174 GCTGAAGAGACCAAGAAATGGTGTACAGCAGTCTCCAGTTAACCGACTCCCCAGACACC 3233
QY 3015 ACAGAGAAAGCCACCCAGTTCAGAGGTAGAGAGTGGTGTGTAGATACAGAGAAGAG 3074
DB 3234 ACAGAGAGGCCACTCCGGTGCAGAGGTGGAAGGTGGCGTACCTGCATAGAGACAA 3293
QY 3075 GAGCGCCAGACGCGAGGCCATCTCCAGCCGTTGACAGAAAGTGAAGAGGAGTCCAG 3134
DB 3294 GAGAGGGGACTCAAGAGGTCTCCAGGAGTGGCAGAAAGTGAAGAGGAATCCAG 3353
QY 3135 GTGCTCGCAACC---CAGACTGTGCAGAGAACGGGGTCAAAAGCACTGGAGAGGTTGAG 3191
DB 3354 CTGCCTGGCCCGGTGGCCAGAGATGTGCTTCAGCCTGTGCGAGAGCAGAGCCAGAA 3413
QY 3192 GAGGTAGAGGAGGACTCCGAAGTCTGGCTTCGAGAGAAAGAGAGGAGCTTATGCCGAAA 3251
DB 3414 AGACCAAGAGAGGAGGCTGAAGCGTCCGGTCTGAAGAAAGAGACGAGTGTGTTGAAA 3473
QY 3252 GGACCCGTGAGGAGAGTGTGAGCTGAGCATCTTGCACAGGSCCTCTGAGACTGGAAGAGCT 3311
DB 3474 GTAGATGCTCAGGAGGCAAAACTGAGCCCTTTTACAGAGGAAAGGTGGTGGGCGACACC 3533

QY 3312 ACTCCAGAGAGCCCTTG---AAGTTCTCTGAAGTCAAGCAGATGTAGAC-----CAT 3359
DB 3534 ACCCCAGAAAGCTTTGAAAAAGCTCTCTCAAGTCACAGAGAGCATAGAGTCAAGTGAGCTT 3593
QY 3360 GTCGCCACAGTGCCA-----GGTTATCAAGCTCCAGCAGCTGATGGAA 3401
DB 3594 GTAAACCATTTGTCAAGCGNAACTTAGCTGGGGTAAATCAAGAGATGGTATGGAA 3653
QY 3402 CAGGCCGTGGCCCTCGAGTCATCCGAAACCTTGACAGACAGTGAAGCAAAATGGAAGCACT 3461
DB 3654 CAGGCTATCCCTCCCTGACTCGGTGGAAACCCCTACAGACAGTGAAGTGAAGCAACC 3713
QY 3462 CCCTTAGCAGATTCAGACACTGCAGATGGGACACAGCAAGATGAACCATTTGACAGCCAG 3521
DB 3714 CCCGTAGCCGACTTTGACGCCACCGGCAACAACCCAGAAAGCAGAGATTTGGAATCCAT 3773
QY 3522 GACAGTAAAGCCACTGCAGCTGTCCAGCTGCAGAGTCAAGGTCAACAGAAAGAGGGCGCTACT 3581
DB 3774 GAGGAGATGA---GGTGCATCTGGTACCAGTCAGGGGCACAGACGAGGAGGTTCTCT 3830
QY 3582 GCTCAGAAAGAGGAGCCTTCGACACTACCTTAATAATGTTCCAGCCCAGGAAGAACATGGG 3641
DB 3831 GCACAGAAAGAGAGGCTCCAGCACCTTCAGTTTGTGTTCCAGGAAGAACTAAAGAA 3890
QY 3642 GAAGAACCGAAGAGATGTTCTTGAAACCTACAGCAAGAGCTTACTGCTGACGCCGTG 3701
DB 3891 CAATCAAGATGGAAGACACTCTAGAGCATACAGATAAAGAGGTGTCACTGGAATCTGTA 3950
QY 3702 CCCGTTCTGCAAGAGCTGAGGTGGGTCAAGAGGGTGA-----GGTTGACTGGTTG 3752
DB 3951 TCCATTCTGTCAAAGACTGAGGGGACTCAAGAGGCTGACCAAGTATGCTGATGAAGAAACC 4010
QY 3753 GATGGAGAAAAAGTCAAGAAAGAACAGGAGTGTGTTGTACACTCTGG-----ACCCAAC 3806
DB 4011 AAGAGCTACCAATTTTTCGAGGACTTGAGGGTCTATAGACACAGGCATAAACAGTCA 4070
QY 3807 AGTCAAAAGGCTGCTGATGTGACATATGACAGTGAAGTATGAGAGTGGCGGTGTGAG 3866
DB 4071 CGGGAAGAGGTCATGAAAGTTGCCCTTAAAGGTGAAGGGACAGAAAGCTGAATGTAAA 4130
QY 3867 GAAAGGAGAGTACTGAAGTGCAGAG-----TCTTAGCCTGGAGGAGGA 3911
DB 4131 AAGATGATGCTCTTGAACCTGCAGAGTCAAGTAAAGTCTCTCACTCCCCCTGGAGAGA 4190
QY 3912 GAGATGGAATACTCAGCTGTGAAAAAGGAGAAAGGAGACAAAGCCAGAGCAAGTGAAGTGA 3971
DB 4191 GAGATGGTAGTTCAGTTCGAAAGGGAGAAACAGAGACAGAGCCCACTGTGAATGAA 4250
QY 3972 GAAGG---TGAGCAGGAAACAGCCGCTCTCTGAGCATGAAGGAAACCTACGGGAAGCCAGTC 4028
DB 4251 GAGAAGCTTGAGCACGAAACAGCTGTTACCGTATCTGAAGAGGTGAGTAAGCAGCTCCTC 4310
QY 4029 CTGACACTTGACATGCCAGCTCAGAGAGGGGGAAGGCATCGGAAGCCTTGGAGGAGAC 4088
DB 4311 CAGACAGTGAATGTGCCCATCATAGATGGGGCAAGGAAAGTCAAGCAGTCTTGGAGGAGAC 4370
QY 4089 CTTCTCTC---TCCAGACCAAGACAAAGCAGGTTGCATAGAGGTTCAAGTTCAAAGCCTG 4145
DB 4371 CTTCTCTCTGCTAGTCAAGAGGAGGCAAGTATGCCAATAATCAAGTTCAGAGCTCT 4430
QY 4146 GACACACAGTCTCACTCAAAACAGCAGAAAGCTGTGGAATA-----GGTCATAGAAACCGTT 4199
DB 4431 GAGGCATCATTTCACTCTAACAGCGGCTGCAGAGGAGGAAAGGCTTTAGGAGAACTGCC 4490
QY 4200 GTGATTTCAGAGACAGGTGAAAGTCCAGAGTGTGTAGGTGCACTATTATTACAGCTGAG 4259
DB 4491 AACATTTTGAAGAACAGGTGAAACGTTGGAGCCTGCGAGGTGCACATTTAGTCTTGGAGAG 4550
QY 4260 AAGTCTCTGCAAGCGGTGGCCACTGGACTCTTCAGCATTCAGAGGACACGGTACCCCTG 4319
DB 4551 AATCTCTGAAAAAATGAAGACTTTGCGCTCATCCAGGGGAAGATGCTGTGCCACA 4610
QY 4320 GGGCCTGAGTCTCAGGAGAGAAATCCATCCCAATCATAGTAATCTCTGCTCTGAAAGCACC 4379

DT 13-FEB-2002 (first entry)
XX DNA encoding novel human diagnostic protein #21009.
DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
KW Homo sapiens.
OS WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US08631.
XX 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
PI P-PSDB; ABG21018.
DR New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity -
PS Claim 1; SEQ ID No 21009; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 6886 BP; 2142 A; 1483 C; 1908 G; 1352 T; 1 other;

Query Match 31.2%; Score 1920.6; DB 23; Length 6886;
Best Local Similarity 67.3%; Pred. No. 0;
Matches 3255; Conservative 0; Mismatches 1334; Indels 244; Gaps 29;

QY 3 GGCGCGGGAGTAGAAGACCCACTGAGCCATGGCGCGAGCGAGTTCACCGAGACGCG 62
DB 162 GGCTAGGCGGGAGAAAGTTCGGAGGAGCCATGGGGCGCGGAGCTCCACCGAGACGCG 221

QY 63 AGCCCGAGCAGCGC---GCGGGAGCGACACCGGAGCGAGTGGTCTCAGTGGCCAT 119
DB 222 AGCCCGAGCAGCGCGCCGAGGGAGGTTCACCGCGGCTGAGCCCGAGCGCCAGCGCGCG 281

QY 120 GGGCGCCGACGTGAAGCCTCGGGA-----GCAGCTGGAGACCCCGCGACGCGGAC 170
DB 282 GGCCCTCGCGGAGCGGGCGCCAGACACACCGGACCCCGCCATCGCTGCTCGGAC 341

QY 171 CCCGCCCAAGCTCCACAGAGAAATGGCCAGGTGTCTTGTCTCAACGCGGTAGCTGAA 230

Db 342 CCCGCCCAAGCTCCCTACAGAGAAATGGTCAGCTGTCCACCATCATGCGGTAGCTGAG 401
QY 231 CAAGGAGATGCTCATGTCTCAAGAGGAAACCCAGGAGG----- 268
DB 402 CAAGATGAGCTCAGCCCTCCAGGAGGGTGACCTAAATGGCCAGAAAGGAGCCCTGAAACGGT 461
QY 269 -----GCAGGAGGAAAGAGTCTGTGATGAGGATGTTGGACACGCA 308
DB 462 CAAGGAGCCCTAAACAGCCAGGAGGAAGAAGTCAATGTCCGGCAGGTTGATCACATA 521
QY 309 GAGTCAGAGATGTGAGAGAAAAGACCGAGTTTCAAGAAATGGCGGCAACTCCACAGCT 369
DB 522 GACTCTGAAGATGTGAGCGACAGAGACTCCGATACAGAGATGACTACTAAGTCAGCGGTA 581
QY 369 GTTGAAGATATCAAAAGGATGGCAGGAGGAGACATCAAGAAATAATTTGAACAGATCCCT 428
DB 582 GTTCAAGCATACAGATGATGGCAGGAGGAGACACCCGATATATTCGAAACAGATTCTCT 641
QY 429 GCTTCAGAAAACAATGTGGAAGAAATGGTACAGCCTGCTGAGTCCCAGGCTAATGATGTT 488
DB 642 TCTTCAGAAAGCAATTTAGAAGAGCTAAACACACCCACTCAGTCCCAGGCTAATGATATT 701
QY 489 GCGTTCAAGAAAGTATTTAAATTTGTTGTTTAAATTCACGGTGAAGAGGATAAAAAT 548
DB 702 GGAATTTAAGAAAGGTGTTTAAAGTTTGTGGCTTTAAATTTCACTGTGAAAAGGATAAGACA 761
QY 549 GAAAAGTCAGATACTGTCCAACTACTCACTGTCAAGAGGATGAAGCGCAAGGGCGAGAA 608
DB 762 GAGNAGCCTGACACTGTCCAGCTACTCACTGTGAAGAAAGATGAAGGGGAGGAGCAGCA 821
QY 609 GCCTGTCTCGAGCTGAGAGACCAACAGGAGCCAGTGTGAGAGACTGCGCTCGGAGAGTCA 668
DB 822 -----GGGCTCGCGACACACAGGACCCAGCCTTGGGCT-----GGAGAAGCA 866
QY 669 GCATCCAAAGAAAGTGAAGTGAAGCAATCCACAGAGAACGAAGACGACCTTGAAGCAA 728
DB 867 GCATCCAAAGAAAGCGAACCCAAACAATCTACAGAGAAACCCGAAGAGACCTTGAAGCGT 926
QY 729 GAACAGAGCAGCAGAGAAA--TCCCTCTTCAAGCCGATCTGATCAAGCGCTCAGGAGAA 787
DB 927 GAGCAAGCCACGCAAAATTTCTCCCAAGCGGAATCTGGCCAA---GCATGGGAGGA 983
QY 788 AGCCAAGATGAAGGAGAGAGAAAACAAGAGAAAGAGCCACCAAGTCCCAGAAATCCCC 847
DB 984 ATGCAAGAGGAGGAGAGAGAGAAAACAAGAAAGAACCTAGCAAGTCTCAGAAATCTCC 1043
QY 848 GAGCAGCCAGTCAACAGTGAAGACATCTTCTTCAAGAAAGTCTTCTCACTCAGCGTTG 907
DB 1044 GACTAGTCCCGTGACCAAGTGAACAGGATCAACCTTTCAAAAAATTTCTTCACTCAAGTTG 1103
QY 908 GGCCTGCTGGCGCAAGAGACCAAGTTCAGAAATCAAAAGAGGATGATCTGGAAGCTGC 967
DB 1104 GGCCTGCTGGCGCAAAAGACCAAGTTCAGGAAGCCGCAAGGAGGATGAAGTGAAGCTTC 1163
QY 968 CGAGAAAGAAAGGAGCAGAGGAGCAAAAAGTAGACGAGAGAAAGAAAGGAAAGACAGA 1027
DB 1164 AGAGAAAGAAAGGAGCAACAGAGCCAGAAAAGTAGACACAGAAAGAGACGAAAGGACAGA 1223
QY 1028 GCAGCCTCGAGGA-----GCAGAGCCCGCAGAGA 1060
DB 1224 GGTTCGCTCCGAGAAACTGACCCCTCCGAGCAAGCCACCCACAGAGCGCGCAGAAAG 1283
QY 1061 CACAGACCCAGGCGAGTGTCTCAGCAGACTACAGAGAGGTGGAGTGCCTTTGGAAGACCA 1120
DB 1284 TGCCCAAGAGCCCGGTTATCAGCTGAATATGAGAAAGTTGAGCTGCCCTCAGAGGAGCA 1343
QY 1121 GGTGTGTGACCTGAGGCGCATCGTCAGAGAGAGTGTGCTCTCTTTGGCAACGGAAGTGT 1180
DB 1344 AGTCAGTGGCTGCAGGGACCTTCTGAGAGAGAAACCTGCTCCGTTTGGCGACAGAGTGT 1403
QY 1181 TGATGAGAAAGATGGAAAGCCCAAGAA---GTTTGTGAGAGGTCCAGTGAAGCCCGT 1237

QY 3358 ATGTCGCACTGTCCTCA-----GGTTATCAAGCTCCAGCAGCTGATGG 3399
Db 3597 TTGTAACCACTTGTCAAGCGAAACCTTAGCTGGGGTAAATCAACAGGATGGTATGG 3656
QY 3400 AACAGGCGGTGGCCCTGAGTCATCCGAAACCTTGCAGACAGTGCAGACAAATGGAAGCA 3459
Db 3657 AACAGGCTATCCCTGCTGACTCGGTGGAACCCCTTACAGACAGTGCAGTGCAGCA 3716
QY 3460 CTCCCTTAGCAGATTACAGACATGTCAGATGGGACACAGCAAGATGAAACCATTTACAGCC 3519
Db 3717 CCCCCGTAGCGCACTTTGAGCGCACCGACCAACCCAGAAAGACGAGATTGTGGAATCC 3776
QY 3520 AGACAGTAAAGCCATGCACTGCTCAGGCGAGTACAGAGTGCAGAGAAAGAGCGGCTA 3579
Db 3777 ATGAGGAATAGAGTGCATCTGGTACCCAGTCCAGGGGACACAGAGCGACGACGCTTC 3836
QY 3580 CTGCTCAGAAAGAGGAGCCTTCGACACTACTCTAATAATGTTCCAGCCCGAGGAACAATG 3639
Db 3837 CTGCACAGAAAGAGAGCCTCCAGCACCTTCCAGTTTGTGTTCCAGGAAGAACTAAG 3896
QY 3640 GGGAAGAACCCAGGAAGATGTTCTTGAACCTACACAGCAAGAGCTTACTGCTGCAGCCG 3699
Db 3897 AACAAATCAAGATGGAAGACACTCTAGAGCATACAGATAAAGAGGTCTCAGTGGAAACTG 3956
QY 3700 TGCCCGTCTCGGCAACACTGAGTGGGTCAAGAGGTGA-----GGTTGACTGGT 3750
Db 3957 TATCCATTCTGTCAAGACTGAGGGGACTCAAGAGGCTGACCAAGTATGCTGATGAGAAA 4016
QY 3751 TGATGAGAAAAAGTCAAGAAAGAACAGGAGGTGTTGTACACTCTGG-----ACCCA 3804
Db 4017 CCAAGAGCTACCAATTTTTCGAGGACTTGAAGGCTTGAAGGCTTATAGACAGAGGATAACAGTCA 4076
QY 3805 ACAGTCAAAAGGCTGCTGATGTGACATATGACATGAAGTGAAGTGGGAGTGGCCGGGTGTC 3864
Db 4077 GTCGGGAAAAGGTCACTGAAGTTGCCCTTAAAGGTGAAGGACAGAGAAGCTGAATGTA 4136
QY 3865 AGGAAAGGAGAGTACTGAAGTGCAGAG-----TCTTAGCTGGAGGAGG 3909
Db 4137 AAAAGGATGATGCTCTTGAACCTGAGAGTCAAGTAACTCTCTCCATCCCGCTGGAGA 4196
QY 3910 GAGAGATGGAACCTGACGTTGAAAGAGAGAAA---AGGGAGACAAAGCCAGAGCAAGTGA 3966
Db 4197 GAGAGATGATGTTCAAGTTCGAAAGGAGGAGATACCAAGTACAGAGCCCAACCATGTGA 4256
QY 3967 GTGAAG---AAGGTGAGCAGGAACACAGCCGCTCTGAGCATGAGGAACCTTACGGGAAGC 4023
Db 4257 ATGAAGAGAAAGCTTGACACGAAACACGCTGTACCGTATCTGAGAGGTCACTAAGCAGC 4316
QY 4024 CAGTCTGACACTTGACATGCCAGCTCAGAGAGGGGGAAGGCACTGGGAAGCCTTGGAG 4083
Db 4317 TCCTCCAGACAGTGAATGTGCCCATCATAGATGGGGCAAGGAAGTCAAGCAGTTTGGNAG 4376
QY 4084 GAAGCCCTTCTC---TCCAGACCAAGCAAGAGCAGGTTGCATAGAGGTTCAAGTTCAA 4140
Db 4377 GAAGCCCTTCTCCTGCTAGTCAAGAGGAGGAGTATGCACCAAAATTCAGATTGAGA 4436
QY 4141 GCCTGGACACAACAGTCACTCAACAGCAGAAAGCTGTGGAAAAGGTCA-----AGAAA 4194
Db 4437 GCTCTGAGGCATCACTCACTCAACAGCGCTGCAGAGGAGGAAAAGTCTTAGGAGAAA 4496
QY 4195 CGGTTGTGATTTACAGACAGGTTGAAAGTCCAGAGTGTGTAGGTGCACACTTATTACAG 4254
Db 4497 CTGCCAAATTTTGAAGAACAGGTGAAACGTTGGAGCCTGCGAGCTGCAATTTAGTCTGG 4556
QY 4255 CTGAGAGTCTCTGCAACGGGTGGCCACTGGACTTTCAGCATGCGAGAGACAGGCTAC 4314
Db 4557 AAGAGAAATCTCTGAAAAAATGAAGACTTTGCGCTCATCCAGGGGAAAGTCTGTGC 4616
QY 4315 CCCTGGGCGCTGAGTCTCAGGACAGATCCCATCCCAATCATAGTAACCTCTGCTCTGAAA 4374
Db 4617 CCACAGGCGCCGACTGTGAGGCAAAATCGACACAGTGTATGATCTGCTACTACCAAGA 4676
QY 4375 GCACCCCTACATCTCTGACCTTCAAGGAGAAATAAGCCGATCCAGAGAGAGCGATCAGAGG 4434

Db 4677 AAGGCTTAAAGTTCGACCTCGAAGGAGAGAGAAAACCATCATCACTGAAGTGAAGTCAGATG 4736
QY 4435 AAGAGGACACAGCCAGATGCTGGTCTCTGATGCTCAGCGCAAGGAGAGTACAGCAATCGAAA 4494
Db 4737 AAGTCGATGAGCAGGTTGCTTGGC-----AGGAGGTCAAGGTGAGTGTAGCAATTGAGG 4790
QY 4495 AAGTCCTCAAGGCTGAACCTTGAGATCCTGGAACCTTGAGAGTAAAGCAACAAGA-TTGTG 4553
Db 4791 AGGATTTAGAGCCTGAAAATGGGATTTTGGAACTTGAGACCAAAAGCAGTAAAACTTGT 4850
QY 4554 CTGAAGCTCAATTACAGACAGCCGTTGACCG-TTCGCAAGTAC---AGAAACAGCCCCGA 4609
Db 4851 CAAAACATCATCCAGACAGCCGTTGACCAAGTTTGTACGTACAGAAAGAAACAGCCACCGA 4910
QY 4610 AACTCATGCTTATGATTTCACAGACCCAGGTTCC 4642
Db 4911 AATGTTGACGCTGAGTTACAGACACAGGCTTC 4943

RESULT 9

ABK6316/C

ID ABK63316 standard; cDNA; 563 BP.

XX AC ABK63316;

XX DT 18-JUN-2002 (first entry)

XX DE Rat sequence differentially expressed in response to a hepatotoxin #1223.

XX DE Rat; ss; hepatotoxin; expressed sequence tag; EST; drug screening;

XX DE differential expression; centrilobular necrosis; steatosis.

XX OS Rattus norvegicus.

XX PN WO200210453-A2.

XX PD 07-FEB-2002.

XX PF 30-JUL-2001; 2001WO-US23872.

XX PR 31-JUL-2000; 2000US-222040P.

XX PR 02-NOV-2000; 2000US-244880P.

XX PR 11-MAY-2001; 2001US-290029P.

XX PR 15-MAY-2001; 2001US-290645P.

XX PR 22-MAY-2001; 2001US-292336P.

XX PR 06-JUN-2001; 2001US-295798P.

XX PR 13-JUN-2001; 2001US-297457P.

XX PR 19-JUN-2001; 2001US-298884P.

XX PR 09-JUL-2001; 2001US-303459P.

XX PA (GENE-) GENE LOGIC INC.

XX PI Mendrick D, Porter MW, Johnson KR, Castle AL, Elashoff MR;

XX XX WPI; 2002-241625/29.

XX XX Predicting toxic effects of compounds or the progression of these toxic

XX XX effects by determining the changes in gene expression in tissues or

XX XX cells exposed to the toxin and comparing these to gene expression in

XX XX unexposed tissues or cells -

XX XX Claim 1; Seq ID No 1223; 239pp; English.

XX XX The invention relates to methods for predicting toxic effects of

XX XX compounds or the progression of these toxic effects by determining the

XX XX global changes in gene expression in tissues or cells exposed to the

XX XX toxin and comparing these to gene expression in unexposed tissues or

XX XX cells. Also included are methods of predicting at least one toxic

XX XX effect of a compound or progression of a toxic effect, preferably the

XX XX hepatotoxicity of a compound, comprising detecting the level of

XX XX expression in a tissue or cell sample exposed to the compound of two or

XX XX more genes listed in the specification, where differential expression of

CC The genes is indicative of at least one toxic effect or progression.
CC The method can also be used to identify an agent which modulates the
CC toxic response and predict cellular pathways that a compound modulates
CC in a cell. The methods utilize a set of at least two probes (on a solid
CC support in kit form), where each of the probes comprises a sequence that
CC specifically hybridizes to a gene listed in the specification, a computer
CC system comprising a database containing information identifying the
CC expression level in a tissue or cell sample exposed to a hepatotoxin of a
CC set of genes comprising at least two genes listed in the specification,
CC and a user interface to view the information used to present information
CC identifying the expression level in a tissue or cell of at least one gene
CC listed in the specification. The method is useful for elucidating global
CC changes in gene expression and for identifying toxicity markers in
CC tissues or cell exposed to a known toxin. The genes may be used as
CC toxicity markers in drug screening and toxicity assays. The genes and
CC gene expression information may be used as diagnostic markers for the
CC prediction or identification of the physiological state of tissue or cell
CC sample that has been exposed to a compound or agent. Hepatotoxicity
CC is characterised by centrilobular necrosis and steatosis. The present
CC sequence is an expressed sequence tag (EST) or cDNA derived from a gene
CC which is differentially expressed in response to a hepatotoxic agent.

Sequence 563 BP; 191 A; 107 C; 110 G; 155 T; 0 other;

Query Match	8.9%	Score 550	DB 24	Length 563
Best Local Similarity	99.1%	Pred. No. 8e-132		
Matches 553, Conservative	0	Mismatches 5	Indels 0	Gaps 0

OY	5601	ATTGGAGCGCTTTGAAAAAGCTCCAAAGCTCCAACTGACCTCAGCAAAAGATTAACCT	5560
Db	563	AAAGGAGCGCTTTGAAAAAGCTCCAAAGCTTCCTAACCTTGACCTGACCAACAAATACCT	504
OY	5661	CCTGGCAAGAGAGACAAGCTCTTTTAAAGTTTACGTAGCTGATCTGTGGCTCT	5720
Db	503	CCTGGCAAGAAABACAAGCTCTTTTAAAGTTTACGTAGCTGATCTGTGGCTCT	444
OY	5721	AGTCTCTGAAAAGTGTTGTTTCTATGACACGCGAGCTCAGAAATATAAACCCTATT	5780
Db	443	AGTCTCTGAAAAGTGTTGTTTCTATGACACGCGAGCTCAGAAATATAAACCCTATT	384
OY	5781	TGAAAACATCCGAGATGTCCTCAATTCATCCATGATTTTTTCCCCCTTTTGCTATCCAG	5840
Db	383	TGAAAACATCCGAGATGTCCTCAATTCATCCATGATTTTTTCCCCCTTTTGCTATCCAG	324
OY	5841	TCCAGGTGGAAAAAGCTCTCTCTGTGCAGATTAAAGCCTGTCTCTTAATGATATGGA	5900
Db	323	TCCAGGTGGAAAAAGCTCTCTCTGTGCAGATTAAAGCCTGTCTCTTAATGATATGGA	264
OY	5901	CAATGAGTGTGCTTAAGGCCATGAGATGTTTCTTAATGACAGAAAGATCTGTGTAGCT	5960
Db	263	CAATGAGTGTGCTTAAGGCCATGAGATGTTTCTTAATGACAGAAAGATCTGTGTAGCT	204
OY	5961	TTTTTGTATGTACTCTTCTATGCTGGACCGAATTCATATGACATCGAAGTAGTCCCTG	6020
Db	203	TTTTTGTATGTACTCTTCTATGCTGGACCGAATTCATATGACATCGAAGTAGTCCCTG	144
OY	6021	TTCTTTACAGATGATTTTGTATGATACATCGAGTTGTGTCTGTATTATCTGSCCCT	6080
Db	143	TTCTTTACAGATGATTTTGTATGATACATCGAGTTGTGTCTGTATTATCTGSCCCT	84
OY	6081	TCCTTTAAGAACATGTGCAATTATGTCTTTGGATPAATTTGATTTGACAACTGATTT	6140
Db	83	TCCTTTAAGAACATGTGCAATTATGTCTTTGGATPAATTTGATTTGACAACTGATTT	24
OY	6141	AAATTAACATATTGACT	6158
Db	23	AAATTAACATATTGACT	6

RESULT 10
AAx40328
ID AAx40328 standard; cDNA; 377 BP

AC AAX40328

DT 16-JUN-1999 (first entry)

DE Human secreted protein 5' EST SEQ ID NO:115.

KM Human; secreted protein; EST; expressed sequence tag; diagnosis;
KM forensic; gene therapy; chromosome mapping; signal peptide;
KM upstream regulatory sequence; cytokine activity; cell proliferation;
KM differentiation; hematopoiesis regulation; tissue growth regulation;
KM reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KM thrombolytic; anti-inflammatory; tumour inhibition; ds.

OS Homo sapiens

PN WO9906439-A2

PD 11-FEB-1999
VY

PF	31-JUL-1998;	98WO-IB01233.
YY		

PR 01-AUG-1997; 97US-0904468.
XX

PA (GEST) GENSEI
XX

P1 Lucile A, Dumas Mine Edwards C, Lacroix B
XX

DR P-PSDB; AAY11610.

PT New nucleic acids

PT thyroid and pancreas tissue

PS Claim 1; Page 228-229; 398

CC AAX40251 to AAX40397 represent

CC AY11679, respectively. The proteins g
CC and an N-terminal fragment of a secret

CC sequences can be used for producing secreted human gene products. They
CC can also be used to develop products for diagnosis and therapy. The
CC proteins obtained may have cytokine activity, cell
CC proliferation/differentiation activity, haematopoiesis regulating
CC activity, tissue growth regulating activity, reproductive hormone
CC regulating activity, chemotactic/chemokinetic activity, haemostatic and
CC thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, tumour inhibition activity or other activities. The products
CC can be used in forensic, gene therapy and chromosome mapping procedures.
CC The sequences can also be used for obtaining corresponding promoter
CC sequences. The nucleic acids encoding the signal peptide can be used for
CC directing extracellular secretion of a polypeptide or the insertion of a
CC polypeptide into a membrane, or importing a polypeptide into a cell.

SQ Sequence 377 BP; 111 A; 85 C; 113 G; 66 T; 2 other;

Query Match	2.8%	Score 173.2; DB 20;	length 377;
Best Local Similarity	68.9%	Pred. No. 4; 3e-34;	
Matches 253; Conservative	1;	Mismatches 104;	Indels 9; Gaps 1;

2455 AAGCCACTGTGGAAGACTCAGGGCCAGTGGAGATAATGAGGACGACCCTAATGTCCAG 2514

Dd 1 AAGCCCTGTTGAGACGCAGGGCCACACAGGGGCCAACGAGATGACTCTGATGTCCCG 60

2515 CCGTCGTGCCCTCTGTCTGAGTATAATGCAGTGGAGAGGAGAGAATCG-----AAG 2563

61 CCGTGGTCCCTCTGTCTGAGTATGATGCTGTAGAAAGGAGAAATGGAGGCACAGCAAG 12

Dy	2566	CCGAGGGGAATACGAGCTGCACCCTGCTGGGGGGCTGTGTACGTGCCAGGAGCTCA	2625
Dh	121	CCCCAAAAGCGCCAGAGCCTTCCGAGAGAAAGAGAGCCACTMGAGGTGTCTSAGAGCTCA	180

Db 121 CCCAAAAAGCCAGAGCAGCCCGAGCAGMAGCCAGCCACCTGAGGTGTCKSAGAGACTCA 180
 Ov 2626 GTAGACTCTGTGCCCACTGTGAGTGTCCGAGTCATTATGSGGACCAAGGAGTCAACCA 2683

QY 2626 GTAGACTCTGGTCCACACTGTGAGTGTCCGAGTCATTGATGGGACCAGGGCAGTCACCA 2685

Db 181 GCGAGAGTCAGTTCATATGATGCGACAGCTGCGCTGACGGACGAGGCGACTACCA 240
 Qy 2886 GTGTCGAAGACGGTCTCTTCGTTGGATATCGGTTCCGTTACAGAACCTCTTGAACACA 2745
 Db 241 TTATTGAAGAAGGCTCTCTTCGTTGGATATCGGTTCCGTTACAGAACCTCTTGAACAA 300
 Qy 2746 CAGCGGAGAGCCACCTGTTGAGAGTCACTGAAAAAGACATCTTCAGAGAG 2805
 Db 301 TAGAAGCTGAAGCCGCACTGTTAACTGAGGAGGTATTGGAAGAAGAGTAATTGCAAG 360
 Qy 2806 AAACCTCC 2812
 Db 361 AAGAAC 367
 RESULT 11
 ABK63568
 ID ABK63568 standard; cDNA; 178 BP.
 AC ABK63568;
 XX
 DT 18-JUN-2002 (first entry)
 XX
 DE Rat sequence differentially expressed in response to a hepatotoxin #1475.
 DE XX
 DE Rat; ss; hepatotoxin; expressed sequence tag; EST; drug screening;
 KW differential expression; centrilobular necrosis; steatosis.
 KW
 XX Rattus norvegicus.
 OS
 XX WO200210453-A2.
 PN
 XX
 PD 07-FEB-2002.
 XX
 PF 30-JUL-2001; 2001WO-US23872.
 XX
 PR 31-JUL-2000; 2000US-222040P.
 PR 02-NOV-2000; 2000US-244880P.
 PR 11-MAY-2001; 2001US-230029P.
 PR 15-MAY-2001; 2001US-290645P.
 PR 22-MAY-2001; 2001US-292336P.
 PR 06-JUN-2001; 2001US-295798P.
 PR 13-JUN-2001; 2001US-237457P.
 PR 19-JUN-2001; 2001US-298884P.
 PR 09-JUL-2001; 2001US-303459P.
 XX
 PA (GENE-) GENE LOGIC INC.
 XX
 PI Mendrick D, Porter MW, Johnson KR, Castle AL, Elashoff MR;
 XX
 XX WPI; 2002-241625/29.
 XX
 PT Predicting toxic effects of compounds or the progression of these toxic
 PT effects by determining the changes in gene expression in tissues or
 PT cells exposed to the toxin and comparing these to gene expression in
 PT unexposed tissues or cells -
 XX
 PS Claim 1; Seq ID No 1475; 239pp; English.
 XX
 CC The invention relates to methods for predicting toxic effects of
 CC compounds or the progression of these toxic effects by determining the
 CC global changes in gene expression in tissues or cells exposed to the
 CC toxin and comparing these to gene expression in unexposed tissues or
 CC cells. Also included are methods of predicting at least one toxic
 CC effect of a compound or progression of a toxic effect, preferably the
 CC hepatotoxicity of a compound, comprising detecting the level of
 CC expression in a tissue or cell sample exposed to the compound of two or
 CC more genes listed in the specification, where differential expression of
 CC the genes is indicative of at least one toxic effect or progression.
 CC The method can also be used to identify an agent which modulates the
 CC toxic response and predict cellular pathways that a compound modulates
 CC in a cell. The methods utilise a set of at least two probes (on a solid

CC support in kit form), where each of the probes comprises a sequence that
 CC specifically hybridises to a gene listed in the specification, a computer
 CC system comprising a database containing information identifying the
 CC expression level in a tissue or cell sample exposed to a hepatotoxin of a
 CC set of genes comprising at least two genes listed in the specification,
 CC and a user interface to view the information used to present information
 CC identifying the expression level in a tissue or cell of at least one gene
 CC listed in the specification. The method is useful for elucidating global
 CC changes in gene expression and for identifying toxicity markers in
 CC tissues or cell exposed to a known toxin. The genes may be used as
 CC toxicity markers in drug screening and toxicity assays. The genes and
 CC gene expression information may be used as diagnostic markers for the
 CC prediction or identification of the physiological state of tissue or cell
 CC sample that has been exposed to a compound or agent. Hepatotoxicity
 CC is characterised by centrilobular necrosis and steatosis. The present
 CC sequence is an expressed sequence tag (EST) or cDNA derived from a gene
 CC which is differentially expressed in response to a hepatotoxic agent.
 XX
 SQ Sequence 178 BP; 41 A; 30 C; 33 G; 74 T; 0 other;
 Query Match 2.6%; Score 162.2; DB 24; Length 178;
 Best Local Similarity 97.8%; Pred. No. 1.1e-31;
 Matches 175; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
 Qy 5962 TTTTGTGATTCTCTCTATGCTGGACCGAATTCATATGCAGATCGAAGTACTCTGT 6021
 Db 1 TTTTGTGATTCTCTCTATGCTGGACCGAATTCATATGCAGATCGAAGTACTCTGT 60
 Qy 6022 TCTTTACAGATGGTATTTTGTGATAGTACTGAGTCTCTCTGTATATCTGTCCCTT 6081
 Db 61 TCTTTACAGATGGTATTTTGTGATAGTACTGAGTCTCTCTGTATATCTGT-CCCTT 119
 Qy 6082 CTTTAAAGAACATGTTGCAATTATGTTCTTTGGATAAATGTTGATTTGACAACCTGATT 6140
 Db 120 CTTTAAAGAACATGTTGCAATTATGTTCTTTGGATAAATGTTGATTTGACAACCTGATT 178
 RESULT 12
 AAV86620
 ID AAV86620 standard; cDNA; 483 BP.
 XX
 AC AAV86620;
 XX
 DT 27-APR-1999 (first entry)
 XX
 DE EST clone BP146.
 XX
 KW Expressed sequence tag; secreted protein; haematopoiesis regulator;
 KW tissue growth; activin; inhibitor; tumour invasion suppressor; EST; human;
 KW chemotaxis; chemokinesis; haemostasis; gene therapy; thrombolysis;
 KW receptor; ligand; anti-inflammatory; tumour inhibitor; ds.
 XX
 OS Homo sapiens.
 XX
 XX WO9845435-A2.
 PN
 XX 15-OCT-1998.
 PD
 XX 10-APR-1998; 98WO-US06954.
 PF
 XX 10-APR-1997; 97US-0835913.
 PR
 XX (GEM) GENETICS INST INC.
 PA
 PI Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;
 PI Racie LA, Spaulding V, Treacy M;
 XX
 XX WPI; 1999-070076/06.
 XX
 PT New polynucleotides encoding human secreted proteins - derived from
 PT e.g. human blood, kidney, foetal lung, placenta, testes, brain,
 PT ovary, pituitary, retina and colon cDNA libraries
 XX

PS Claim 1; Page 294; 633bp; English.

XX This sequence represents an expressed sequence tag (EST), and is a
CC polynucleotide of the invention. The polynucleotides of the invention are
CC all secreted EST sequences isolated from a variety of human tissue
CC sources. The EST sequences and proteins encoded by them are predicted to
CC have useful biological activities which would make them suitable for
CC treating, preventing or ameliorating medical conditions in humans and
CC animals, although no supporting data is given. Suggested activities
CC include nutritional activity, immune stimulating or suppressing activity,
CC hematopoiesis regulating activity, tissue growth activity,
CC activin/inhibin activity, chemorectic/chemokinetic activity, haemostatic
CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, cachectin/tumour invasion suppressor activity, tumour inhibition
CC activity. The EST sequences are also stated to be useful for gene
CC therapy.

XX Sequence 483 BP; 156 A; 103 C; 133 G; 90 T; 1 other;

Query Match 1.7%; Score 105.2; DB 20; Length 483;

Best Local Similarity 60.4%; Pred. No. 1.1e-16;

Matches 236; Conservative 0; Mismatches 134; Indels 21; Gaps 3;

```

OY 3811 AAAAGCTGCTGATGTCATATGTCAGTGAAGTGGAGTGGCGGTGCGAGAA 3870
    |||||
DB 44 AAAAGTCTCTGAAGTTCCTTAAAGTGAAGGACAGAGAACTGAATGTAAGG 103
    |||||
OY 3871 AGGAGTCTCTGAAGTTCAGAG-----TCTTACCTGAGAGGAGAGAGA 3915
    |||||
DB 104 ATGAGTCTCTTGAAGTTCAGAGTCAAGTCAAGTCTCTTCAATCCCGGAGAGAGAGA 163
    |||||
OY 3916 TGGAAATCTGACGTTGAAAAAGGAAAAAGGAGACAAAGCCAGAGCAAGTGAAGAG 3975
    |||||
DB 164 TGGTGTCTCAAGTCCGAAAGGAGGAGAAACAGAGAGCAAGCCATGTGAATGAAGGA 223
    |||||
OY 3976 G---TGACAGAGAAACAGCGCGCTCGAGCATGAAGGAACCTACGGGAAAGCAGTCTGA 4032
    |||||
DB 224 AGCTTGACACACAAACAGTGTACCGAATCTGAAGAGTCAAGTCAAGCAGTCTTCCAGA 283
    |||||
OY 4033 CACTTGACATGCCCACTCAGAGAGGGGAGAGCACTGGGAAGCCTTGGAGAGAACCCCTT 4092
    |||||
DB 284 CAGTGAATGTGCCCATCATAGATGGGGCAAGAGATCGAGAGTTTGGAGAGAACCCCTC 343
    |||||
OY 4093 CTC---TCCAGAGCCAGACAAAGCAGTTGCAATGAGAGTTCAAGTCAAAGCCTGGACA 4149
    |||||
DB 344 CTCCTGCTAGTCAAGAGGAGGAGATGACCAAAATTCAAGTTCAAGAGCTCTGAGG 403
    |||||
OY 4150 CAACGATCACTCAAAACAGAGAAAGCTGTGGA 4180
    |||||
DB 404 CATCATTCATCTTAACAGCGGCTGAGAGGA 434
    |||||

```

RESULT 13

ABV56830/c

ID ABV56830 standard; cDNA; 459-BP.

XX ABV56830;

XX 17-SEP-2002 (first entry)

XX Human prostate expression marker cDNA 56821.

XX Human prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

XX pharmacogenomic marker; gene; ss.

XX Homo sapiens.

XX WO200160860-A2.

XX 23-AUG-2001.

XX 20-FEB-2001; 2001WO-US05171.

XX

PR 17-FEB-2000; 2000US-183319P.
PR 16-MAR-2000; 2000US-189862P.
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege WO, Monahan JE,
PI WPI; 2001-662795/76.
XX
XX

PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer
XX
XX
XX Claim 1; Page 10950; 11750bp; English.

CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
XX

Sequence 459 BP; 143 A; 99 C; 59 G; 158 T; 0 other;

Query Match 1.7%; Score 101.8; DB 23; Length 459;

Best Local Similarity 70.5%; Pred. No. 8.3e-16;

Matches 136; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

```

OY 5963 TTTGATGTGTCTCTCTATCTGACCGAATTCATATGACATGCAAGTGAAGTCTGTT 6022
    |||||
DB 235 TTTGATTTTACTTTTAAAGTGGCGGCATTTACACAGAGATGAATTAAGTCAGTT 176
    |||||
OY 6023 CTTTACAGATGTAATTTGATGATGACTGAGTGTGCTGTGTTATATCTGTGCCCTTC 6082
    |||||
DB 175 TTTTACAAAAGGATTTGAAGAAAACTGAGATGGGTTGGGCCAAAATTGGGCTTTT 116
    |||||
OY 6083 TTTTAAAGCAATGTTGATTTGTTCTTGGATTAATTTGATTTGACATGATTTAA 6142
    |||||
DB 115 TTTTAAAGCAAAAGTGGCCACCATTCATTTGGATTAAGTTGATTTGACATGATTTAA 56
    |||||
OY 6143 ATTAACATATTTG 6155
    |||||
DB 55 AAAAAAAATTGG 43
    |||||

```

RESULT 14

AA186887

ID AA186887 standard; cDNA; 400-BP.

XX AA186887;

XX 06-NOV-2001 (first entry)

XX Human polynucleotide SEQ ID NO 6947.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX vaccine; peptide therapy; stem cell growth factor; hematopoiesis;
XX tissue growth factor; immunomodulatory; cancer; leukemia;
XX nervous system disorders; arthritis; inflammation; ss.

XX

Db 192 ATCAAAGATGGAAGAGCACTTAGAGCATACAGATTAAAGAGGTGTCAGTGAAGAACTGTATC 251
Qy 3704 CGTCTGGCAAGAAGACTGAGTGGGTCAAGAGGGTGA 3739
Db 252 CATCTGTCAAGAGACTGAGGGGACTCAAGAGGCTGA 287

Search completed: December 12, 2002, 16:39:17
Job time : 1217 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 12, 2002, 14:13:30 ; Search time 166 Seconds
(without alignments)
11380.294 Million cell updates/sec

Title: US-09-902-432-3
Perfect score: 6160
Sequence: 1 cggcgccggaggtagaag.....aaataaacatttgactac 6160

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
- 2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
- 3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
- 4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
- 5: /cgn2_6/ptodata/2/ina/PCTUS COMB.seq.*
- 6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4831.4	78.4	5134	2	US-08-635-121-1
2	2118.8	34.4	6605	1	US-08-769-309A-4
3	2118.8	34.4	6605	3	US-08-394-570-4
C 4	99.8	1.6	7218	1	US-08-232-463-14
C 5	71.2	1.2	1926	4	US-09-249-585A-4
C 6	71.2	1.2	1931	2	US-09-130-114-2
7	70.8	1.1	1926	4	US-09-249-585A-2
8	70.8	1.1	2580	3	US-09-050-863-2
9	70.8	1.1	2580	4	US-09-359-081-2
C 10	70.8	1.1	5452	2	US-09-130-114-1
11	70.8	1.1	9600	4	US-08-910-647-1
12	70.8	1.1	9600	4	US-09-620-925-1
13	70.8	1.1	10596	1	US-07-884-811-15
14	70.8	1.1	10596	1	US-07-885-971-15
15	70.8	1.1	10596	1	US-08-087-783A-15
16	70.8	1.1	10596	1	US-08-194-088B-15
17	70.8	1.1	10596	1	US-08-194-087-15
18	70.8	1.1	10596	5	PCT-US93-04648-15
19	66.4	1.1	3489	2	US-08-728-323A-1
20	66.4	1.1	3489	4	US-09-298-568-1
C 21	66.4	1.1	32207	2	US-08-770-379-20
C 22	66.4	1.1	32207	4	US-08-757-669A-20
C 23	66.4	1.1	32207	4	US-09-230-371A-20
C 24	64.6	1.0	16442	3	US-08-781-891-208
25	63.4	1.0	7218	1	US-08-232-463-14
26	60.4	1.0	2277	1	US-08-676-967-2
27	60.4	1.0	2277	1	US-08-676-967-2

28	60.4	1.0	2277	2	US-09-098-487-2	Sequence 2, Appli
29	60.2	1.0	3337	1	US-08-072-610-1	Sequence 1, Appli
30	60.2	1.0	3337	2	US-08-719-822B-1	Sequence 1, Appli
31	60.2	1.0	3337	4	US-09-092-458-1	Sequence 1, Appli
C 32	58.8	1.0	51259	3	US-08-781-891-209	Sequence 209, App
33	56.4	0.9	3211	2	US-08-574-959A-8	Sequence 8, Appli
34	56.4	0.9	3211	4	US-09-357-014-8	Sequence 8, Appli
35	56.4	0.9	3901	2	US-08-574-959A-6	Sequence 6, Appli
36	56.4	0.9	3901	4	US-09-357-014-6	Sequence 6, Appli
37	55	0.9	1052	1	US-08-466-603-1	Sequence 1, Appli
38	55	0.9	1052	1	US-08-314-503A-1	Sequence 1, Appli
39	55	0.9	1052	1	US-08-468-066-1	Sequence 1, Appli
40	55	0.9	1052	2	US-08-466-717-1	Sequence 1, Appli
41	55	0.9	1052	3	US-08-466-743-1	Sequence 1, Appli
42	55	0.9	1052	5	PCT-US95-12414-1	Sequence 1, Appli
43	52.6	0.9	1505	1	US-07-915-246-1	Sequence 1, Appli
44	51.8	0.8	43795	3	US-08-742-185-101	Sequence 101, App
45	47.8	0.8	2518	4	US-09-433-699-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-08-635-121-1
; Sequence 1, Application US/08635121
; Patent No. 5910442
; GENERAL INFORMATION:
; APPLICANT: Gelman, Irwin H.
; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10112-0228
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/635,121
; FILING DATE: 19-APRIL-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Richard S
; REGISTRATION NUMBER: 26,154
; REFERENCE/DOCKET NUMBER: A30558 - 165/33603
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-408-2558
; TELEFAX: 212-765-2519
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5134 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; US-08-635-121-1
Query Match 78.4%; Score 4831.4; DB 2; Length 5134;
Best Local Similarity 97.9%; Pred. No. 0;

Db 4300 TTGATGCCGTCGCTGATTTTCAACATTTCTGTTTCAAGACGCTTTGATTTGCCCC 4359
 QY 5386 TTGATGCCGTCGCTGATTTTCAAGACGCTTTGATTTGCCCC 5445
 Db 4360 TTGATGCCGTCGCTGATTTTCAAGACGCTTTGATTTGCCCC 4419
 QY 5446 GGCATACCTAGTTCACCTCTCAAACTGGAGACCTCTTATGATTTATGATATG 5505
 Db 4420 GGCATACCTAGTTCACCTCTCAAACTGGAGACCTCTTATGATTTATGATATG 4479
 QY 5506 TTTATGATGCTCTCTCTCTGATCTGATTTGATTTTCTTAAAGCTTTAAGCAAGCT 5565
 Db 4480 TTTATGATGCTCTCTCTCTGATCTGATTTGATTTTCTTAAAGCTTTAAGCAAGCT 4539
 QY 5566 TTTTGTATTTATGCAATTTATTAAGGGTGTGCAAGCTTAAAGCTTTCCAA 5625
 Db 4540 TTTTGTATTTATGCAATTTATTAAGGGTGTGCAAGCTTAAAGCTTTCCAA 4599
 QY 5626 GGCCTCACTGTAACCTGCGCAAAAGATTAATTCCTGCGCAAGAAAGCAAGCTTTT 5685
 Db 4600 GGCCTCACTGTAACCTGCGCAAAAGATTAATTCCTGCGCAAGAAAGCAAGCTTTT 4659
 QY 5686 TTTAAAGTTTACTGATGCTTGAATCTGTGCGCTTCTAGTCTCTGAAAAGTGTGTTTCC 5745
 Db 4660 TTTAAAGTTTACTGATGCTTGAATCTGTGCGCTTCTAGTCTCTGAAAAGTGTGTTTCC 4719
 QY 5746 TATGACACGAGCTGCAAAATTAATAAATTTGAAATCTCCAGATGTCCTCAATAT 5805
 Db 4720 TATGACACGAGCTGCAAAATTAATAAATTTGAAATCTCCAGATGTCCTCAATAT 4779
 QY 5806 TACCAATGATTTTCTCCCTTTTCTTATGATGATGCAATTAAGTGTGCTTAAAGCTTCT 5865
 Db 4780 TACCAATGATTTTCTCCCTTTTCTTATGATGATGCAATTAAGTGTGCTTAAAGCTTCT 4839
 QY 5866 GTGTGATTAAGCCCTGCTCTTATGATGATGCAATTAAGTGTGCTTAAAGCTTCT 5925
 Db 4840 GTGTGATTAAGCCCTGCTCTTATGATGATGCAATTAAGTGTGCTTAAAGCTTCT 4899
 QY 5926 GATGTTTCTTATGCAAGAAATCTGTGATGATGATGATGATGATGATGATGATGATG 5985
 Db 4900 GATGTTTCTTATGCAAGAAATCTGTGATGATGATGATGATGATGATGATGATGATG 4959
 QY 5986 GATGTTTCTTATGCAAGAAATCTGTGATGATGATGATGATGATGATGATGATGATG 6045
 Db 4960 GATGTTTCTTATGCAAGAAATCTGTGATGATGATGATGATGATGATGATGATGATG 5019
 QY 6046 ATACTGAGTTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 6105
 Db 5020 ATACTGAGTTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 5079
 QY 6106 TTTCTTTGATTAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 6160
 Db 5080 TTTCTTTGATTAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 5134

RESULT 2
 US-08-769-309A-4
 ; Sequence 4, Application US/08769309A
 ; Patent No. 5741890

GENERAL INFORMATION:
 ; APPLICANT: Scott, John D.,
 ; APPLICANT: Naurett, Brian J.,
 ; APPLICANT: Kluack, Theresa M.
 ; TITLE OF INVENTION: Protein Binding Domains of Gravin
 ; NUMBER OF SEQUENCES: 24
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 ; STREET: 6300 Sears Tower/233 South Wacker Drive
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: United States of America
 ; ZIP: 60606-6402
 ; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/769,309A
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: NO. 5741890and, Greta E.
 ; REGISTRATION NUMBER: 35,302
 ; REFERENCE/DOCKET NUMBER: 27866/33451
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 312-474-6300
 ; TELEFAX: 312-474-0448
 ; TELEX: 25-3856
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 6605 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: CDNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 192..5531
 ; US-08-769-309A-4

Query Match 34.4%; Score 2118.8; DB 1; Length 6605;
 Best Local Similarity 64.2%; Pred. No. 0;
 Matches 4206; Conservative 0; Mismatches 1802; Indels 546; Gaps 46;

QY 3 GGGCCGCGGAGTGAAGACCACTGAGCCATGGGCGGAGGAGTTCACGAGGAGG 62
 Db 162 GGGTGGGCGGAGGAGTGAAGACCACTGAGCCATGGGCGGAGGAGTTCACGAGGAGG 221
 QY 63 AGCCCGGAGGAGGCG---GCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 119
 Db 222 AGCCCGGAGGAGGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 281
 QY 120 GGGCCGCGGAGTGAAGACCTGCGGGA-----GCACTGAGAGACCCGCGGAGG 170
 Db 282 GGGCCCTGCGGCGGAGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 341
 QY 171 CCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 230
 Db 342 CCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 401
 QY 231 CAAGGAGATGTCATGTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 268
 Db 402 CAAGGAGATGTCATGTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 461
 QY 269 -----GCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 308
 Db 462 CAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 521
 QY 309 GAGTGAAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 368
 Db 522 GAGTGAAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 581
 QY 369 GTTGAAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 428
 Db 582 GTTGAAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 638
 QY 429 GCTTCAAGAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 488
 Db 639 TCTTCAAGAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 698
 QY 489 GCTTCAAGAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 548
 Db 699 GATTTAAGAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 758
 QY 549 GAAAGTCAAGATGTCACACTCACTGCAAGAGGAGGAGGAGGAGGAGGAGGAGG 608

OY	2718	GCCTCCGTAACAGAA	CCTCTTTGAACACACAGCGGGAGAACCATGCACCTGTGGAAG	2777
Db	2934	GCTTCAGTACAGAA	CTCTTGAACAGTGAAGCTGAAGCCGCACTGTTACTGAAG	2993
OY	2778	GTCACGTAAAAAG	CATCATTCGC---AGAGAACTCTGTGCTCACCAAGCTTACCA	2834
Db	2994	GTATTGGAAAGAA	GTAAATTGCGAAGAAAGAACCCCCACGGTTACTGAACCTCTGCA	3053
OY	2835	GAGGTTAAAGT	GGCCATTAACACATGTGTCCACAGTGAAGTGTGATTTTCACTCAGAAAGCT	2894
Db	3054	GAGAACAGAGAG	GGCCCGGGCGACAGGTGCTTGTGAAGCGAAATTGACCCCGAAGCT	3113
OY	2895	GTGACAGCCACAG	AGACCTCAGAGAGCTTCCTCGTACGAAAGATTACCGAAGCATCGGG	2954
Db	3114	GTGACAGCTGAG	AAATCTGACAGGGCCATTGGTTCCGAAAGAAACCGAAGCATCTGCT	3173
OY	2955	GCCGAAAGAC	CCACAGCATGTGTGTCCGACATTTCCAGCTGACTGATCTCCGACACCC	3014
Db	3174	GCTGAAGAGAC	CAACAGAAATGTGTGACAGCTGTCCCAATTAAACCGACTCCCGACACCC	3233
OY	3015	ACAGAGAAAG	CAACCCAGTTCAAGAGGTGAAGAGGTGTGCTGATACAGAAAGAAAG	3074
Db	3234	ACAGAGAGG	GCACCTCCGGTGCAGAGAGTGAAGGTGGCTACCTGACATGAGAGCATA	3293
OY	3075	GAGCGCCAG	ACGACCATCTTCCAGACCCGTTCAGACMAAGTGAAGAGAGATCCAG	3134
Db	3294	GAGAGCGG	CACTCAAGAGGTCTTCCAGAGCTGTGACAGAAATGTGAAGAGATCCAG	3353
OY	3135	GTGCTGCA	ATCC---CAGACTGTGCAGAGAACGGGGCTCAAAAGCACTGAGAAAGTTGAG	3191
Db	3354	CTGCTGTGC	ACCGGTGGCCAGAAAGATGTGCTTCACTGTGCAGAGAACAGAGGCATAA	3413
OY	3192	GAGGTAGAGAG	AGATCCGAAGTCTGCTCGAGAAAGAGAAAGACCTGTATGCGCAAA	3251
Db	3414	AGACCGAAG	AGACAGCTGAAGCTCGGGCTGAABAAAGAGACGGAATGTATGTTGAA	3473
OY	3252	GCACCTGTC	CAGAACTGAGCTGACATCTTTCACAGGGCTGTGACCTGACACAGCT	3311
Db	3474	GTAGATGCT	CAGAGGCAAAATCTGACCTTTTACCAAGAGMAAGTGTGTGGGCGACAC	3533
OY	3312	ACTCCAGAG	ACCTCTTGA---AAGTTCCTGAAGTCAACGCGACAGATGTGAC-----CAT	3359
Db	3534	ACCCAGAA	ACCTTTGAAGAGCTTCCAGATCAACAGAGCATGAGTCACTGATGACTT	3593
OY	3360	GTCCGAC	AGTCCCA-----GTTATCAAGCTCCAGCACTGTATGAA	3401
Db	3594	GTAAACCA	CTTGTCAAGCCGAAACCTTATAGTGGGGTAAATTCACAGAGATGTGTATGAA	3653
OY	3402	CAGGCGGT	GGCCCTGTGATCATCCGAAACCTTACACAGACAGTGAACAAATGGAAGCACT	3461
Db	3654	CAGGTTAT	CCCCCTGTGCTGGTGAAGACCTTACACAGACGTGAGCTGATGGAAGCACCC	3713
OY	3462	CCCTTACAG	ATTCAGACATCTGCAGATGTGGACACACCAAGATGATAACCATTTACAGCCAG	3521
Db	3714	CCCGTAG	CCGACTTTGACACACAGGACACACCAAGAAAGACGAATGTGGAATTCAT	3773
OY	3522	GACAGTAA	AGCCACTGCACTGTACGGCACTCAACAGTCAACAGAGAAAGCGCGTACT	3581
Db	3774	GAGGAGAT	TGA---GGTGCACTGTGTTCAGAGTCAAGGGGACAGAAAGCAAGGCACTCT	3830
OY	3582	GCTCAGAA	AGAGAGCTTTCGACACTACCTAATTAATGTTCAGGCCACAGGAATAACATGGG	3641
Db	3831	GCACGAA	AGAGGCGCTCCAGCACTTTCAGATTTGTGTTCAGAGAAATACTAAGGA	3890
OY	3642	GAAGAAC	CCAGAAAGATGTCTTGAACCTTACACAGCAAGAGCTTACTGCTCAGCCGTG	3701
Db	3891	CAATAAAG	ATGAAAGACACTCTTGAACATATACATTAAGAGGTGTCAATGGAACCTGTA	3950
OY	3702	CCCGTT	CTTGCAAAAGCTGAGAGTGGGTCAAGAGGTGA-----GTTACATGTGTG	3752
Db	3951	TCCATT	CTGTCTCAAAAGACTGAGAGGGGACTCAAGAGGCTGACCATGATCTGATAGAAAAC	4010

Qy	3753	GATGGAAAAAATCTCAAAAGAAACAGAGGGTGTGTAACCTCTGG-----ACCGAAC	3806
Db	4011	AAAGACTACATTTTTCGAAAGACTTGAAGGGCTTAATAGACAAGCATPACGTCACT	4070
Qy	3807	AGTCAAAGGCTCTCATGTGACATATGACATGTGAAGTGAAGGAGTGGCCGGGTGTCA	3866
Db	4071	CGGAAAAAGGTCACTGAAGTGTCCCTTAAGGTGAAGGACAGAAAGACGTGAATGTAAA	4130
Qy	3867	GA AAAAGAGATCTGAAGTGCAGAG-----TCTTAAGCTTGGAGAGGGA	3911
Db	4131	AAGGATATCTCTTGAATCTGCAGATGCACGGTAAGTCTCTCCATCCCTCGTGAAGGA	4190
Qy	3912	GAGATGGAACCTACGTTGAAAAGGAAAGGAGCAAGGAGCAAGGCCAGAGCAAGTAGTAA	3971
Db	4191	GAGATGTGATTCAAGTCTGAAGGGGAGAAAACAGAACAGAGCCACCCATGTATAAGAA	4250
Qy	3972	GAAGG---TGAGAGGAAACAGCCGCTCTGTGACATGAAGAACCTACGGGAACCACTC	4028
Db	4251	GAGAACTTGAGCACTGAACAGCTGTTTACCGTATCTGAAGAGTCAATGAACACTCTCTC	4310
Qy	4029	CTGACACTTGAATGCCAGCTCAGAGAGGGGAAAGGACCTGGAAAGCTTTGAGAGAGC	4088
Db	4311	CAGACAGTAATGTGCCATCATATAGTGGGGCAAGAGAAAGTCAGACGTTTGAAGAGAAC	4370
Qy	4089	CTTTCTC---TCCACACCAAGCAAGAGGTTGCTATAGAGTTCAAGTTCAAGGCTTG	4145
Db	4371	CTCTCTCTCGCTTAAGGTCAAGAGGAGGACATATCACAAATTCAGATTCAGAGCTCT	4430
Qy	4146	GACACAAACGTCACTCAACAGCAGAAAGCTGTGAAAA-----GGTCAATGAAGCGTT	4199
Db	4431	GAGGATCATTTCACTCTTACACAGGGCTTGACAGAGAGAAAGGCTTTAGAGAAACTGCC	4490
Qy	4200	GTGATTTCAAGACAGGTGAAGTCCAGATGTGTAGTGCACATTAATTAACAGCTGAG	4259
Db	4491	AAACATTTAGAAACAGGTGAAAGTTGGAGCTGCAGAGGTGCACATTTAGTTCTGAAAGAG	4550
Qy	4260	AAATCTCTGCAACGGGTGGCCACTGTGAATCTTTCAGCATGACAGAGACACGGTACCCCTG	4319
Db	4551	AAATCTCTGTAAAAAATGAAAGCTTTGGCCGTCACTCAAGGGGAAAGAGCTGTGTCCACA	4610
Qy	4320	GGGCTGAGTGTGAGCAGAAATCATCCCATCATATGTAACTCTCGCTCTGAAAGCACCC	4379
Db	4611	GGGCCCCATCTCAGGCAAAATTCGACCAAGTATATGATCTGTCTACTTCCAAAGAAAGCC	4670
Qy	4380	CTACATCTCTGACCTTCAAGAGGAAATTAAGCGCATCCAGAGAGAGCGATCAGAGAAAG	4439
Db	4671	TTTAGTTCGACCTGGAAGGAGAAACACATCATCTGAAGTGAATCAGATGAAGTCTC	4730
Qy	4440	GACAAAGCCAGATCTGTGTCTGTATGCTGACGGCAGAGAGAGTACAGCATATGAAAAAGTC	4499
Db	4731	GATGAGCAGGTGTCTGTCC-----AGGAGGTCTAAAGTGAAGTGTAGCAATTTG---AGGAT	4781
Qy	4500	CTCAAGGCTGAACCTGATCTCTGAAACTTGAAGTGAAGCAACAATGTGTCTGAC	4559
Db	4782	TTAGAGCTGAAAATGGGATTTTGGAACTTGAACCCAAAGCATTAACCTGTCCAAAAC	4841
Qy	4560	GTTCATTCAGACAGCCGTTCGACAGTTTCGACAGTAC---AGAAACAGCCCCGAAACTCAT	4616
Db	4842	ATCATCTCAGACAGCCGTTGACAGATTTGTATCGTACAGAAAGAAACAGCCACCGAAATGTGG	4901
Qy	4617	GCTTATGATTCACAGACCAGGTTT-----CTGCATCAGAGCTTTCAG	4660
Db	4902	ACGTCTGATGTTACGACACAAAGTCAAGTATTAAGATGACAGCCAGGACGCTGAGACAG	4961
Qy	4661	CAGGAGCCCAACAGATGCTGACAAAAATGAAGATGCCAAGATGAACAACCCAGTGTCC	4720
Db	4962	GAACAGGAGAAAGAGAGAGAACTTCAGGGCTCTGCAACAGATGAAGAACACCAATTAAT	5021
Qy	4721	GGAGCCCA-----GAGAGACTTGCAGAGCTGACCGTCTTGG	4759
Db	5022	TTAGCCAAAGAGAGATCAGATCAACCGCAGTGGGACAAAGCACTTTCTGATTTTCCAAA	5081
Qy	4760	GGCAGAGGCTCAGGCTCGGAATGCTTGTCCGGCTTGACAGTTGAAGGCGCGGTGTCAA-	4818

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 192..5531
US-08-994-570-4

Query Match 34.4%; Score 2118.8; DB 3; Length 6605;
Best Local Similarity 64.2%; Pred. No. 0;
Matches 4206; Conservative 0; Mismatches 1802; Indels 546; Gaps 46;

QY 3 GGGCGCGGAGTGAAGAGCCATGACCATGGGCGAGGCACTTCCACCGAGAGGG 62
Db 162 GGTAGAGGCGCGGAGAAATGCGAGAGACCATGGGCGCGGAGCTCCACGAGAGCGC 221
QY 63 AGCCCGGAGAGCGCG---GCGGAGAGCGACACCGCGAGAGCTGGTCTCAGTGGCCAT 119
Db 222 AGCCCGGAGAGCGCGCGCGGAGGAGCTCCACGCGGCTGAGCCCGAGCCGAGCGCGC 281
QY 120 GGGCGCGAGCTGAAGCCTCGGGA-----GCAGCTGAAGACCCCGCGAGCGGAC 170
Db 282 GGGCGCTGCGCGAGGCGCGCGACACACCGCGGACCCCGCATGCTGCTCGGAC 341
QY 171 CCGCGCACCAAGCTTCCACAGAGAAATGGCCAGCTCTTCTGTCAACGGCGTAGTTGA 230
Db 342 CCGCGCACCAAGCTTCTCAGAGAAATGGTCACTGTCCACCATCAATGGCGTAGCTAG 401
QY 231 CAAGAGATGTCCATGTCCAGAGAGAAACGAGAGG----- 268
Db 402 CAAGATGAGCTCAGCTCAGAGAGGAGTGAATGCGCAGAAAGAGCCCTGAACGGT 461
QY 269 -----GCAGAGAGAGAAAGTCTGTTGATGAGAGATGTTGAACGCA 308
Db 462 CAAGAGCCCTAAACAGCGAGAGAGAAAGAAATGTCATTGTCACGAGATTGACAGAGA 521
QY 309 GAGTCAGAAATGTGAGAGAAACCGAGTTGAAGAAATGGCGGCACTCCACAGCT 368
Db 522 GACTCTGAAGATGTGAGGAGAAAGAGACTCCGATAAAGATGGCTCAATAGTCACGCGTT 581
QY 369 GTTGAAGATATCAACAAGATGGGAGAGAGAGCATCAAGAAATTAATGAACAGATCCCT 428
Db 582 GTTCAGCATATCAAGATGATGGAGAGAGAG---ACCGAAATATGAAACAGATCTT 638
QY 429 GCTTCAGAAACAATGTGAGAAATGGTACAGCCTGTGATGCCAGCTTAATGATGTT 488
Db 639 TCTTCAGAAAGAAATTTGAGAGCTAAACAACCCACAGATGCCAAGCTTAATGATTT 698
QY 489 GCGTTCAAGAAATGATTTTAAATTTGTTGTTTAAATTCACGCTGAAGAGATTAAT 548
Db 699 GGATTTAAGAAAGTGTTTAAATTTGTTGCTTTAAATTCAGCTGTAAGAAAGATTAAGA 758
QY 549 GAAAGTGAATCTGCTCAACTCACTGTCAGAGAGATGAAGGCGAAGGGGCGAGAA 608
Db 759 GAGAAAGCTGACATGCTCACTGCTCACTGTAAGAAAGATGAAGGAGGAGAGAGAGCA 818
QY 609 GCGTCTGTGAGAGCTGAGACCAACGAGAGCCAGTGTGAGAGACTGCGGTGAGAGTCA 668
Db 819 -----GGGGCTGGGAGCCACAGAGACCCAGCCTTTGGGGCT-----GGAAGAGA 863
QY 669 GCATCCAAAGAAATGAGCTGAAGCAATTCACAGAGAAAGCAAGAGGCACTTGAAGCAA 728
Db 864 GCATCCAAAGAAAGGAGAAACCCAAACAATCTACAGAGAAACCGAAGAGAGCCTTGAAGGCT 923
QY 729 GAAAGAGACAGACAGAAATCCCGCTTCAAGCGAATCTGATCAAGCGGCTGAGAGAGAA 788
Db 924 GAGCAAAACCAAGCAAGAAATTTCTCCCGAGCAATCTGCGCAAGAGCT---GGAAGAGA 980
QY 789 GCAAGAGATGAAGAGAGAGAAACAGAGAAAGAGCCACCAAGTCCCAAGATCCCG 848
Db 981 TGCAAGAGAGAGAGAGAGAGAAACAGAGAAAGAACTTAGAGAAATCTGCAAGATCTCCG 1040
QY 849 AGCAGCCCAAGTCAACAGTGAAGCAACATCTTCTTCAAGAAATGTTCTTCACTCAGGTTGG 908

Db 1041 ACTAGTCCCGTGAACCACTGAAACAGATCAACCTTCAAAAATTTCTTCACTCAAGTTGG 1100
QY 909 GCCGCTGGGCGCAAGAGACCAAGCTTCAAGAAATCAAAAAGAGATGATCTGAAACTGCC 968
Db 1101 GCGGCTGGGCGCAAAAAGACCAAGTTTCAAGAGAGCCGAAGAGATGAAGTGAAGCTTCA 1160
QY 969 GAGAGAGAAAGAGAGCAAGAGGCGAGAAATGACAGAGAGAGAGAGAGAGAGAGAG 1028
Db 1161 GAGAGAGAAAGAGAGCAAGAGCCAGAAATGAGACACAGAGAGAGAGAGAGAGAGAG 1220
QY 1029 CCAAGCTCGAGAG-----GCAGAGAGCGGCGAGAGAC 1061
Db 1221 GTTGCCTTCAGAAATCTGACCCGCTCCGAGACAGCCACCAAGAGAGCGGCGAGAAAGT 1280
QY 1062 ACAGAGAGGCGAGATTGTGACAGACTACAGAGAGAGTGAAGTCTCTTGAAGACAG 1121
Db 1281 GCCAGAGAGCGCGGTTATCAGCTGAATATGAGAAATGAGTGTGCTCCCTCAGAGAGACAA 1340
QY 1122 GTTGTGACCTGAGAGCAATGTCAGAGAGAGAGTGTCTCTTTGGCAACGAAATGTTT 1181
Db 1341 GTCAGTGCTCGCAGAGAGCCTTCTGAAGAGAAACCTGCTCGTTGGCGACAGAGTGT 1400
QY 1182 GATGAGAGATGGAAGGCCCAAGAGAA---GTTGTGAGAGGTCAAGTGAAGACCGTG 1238
Db 1401 GATGAGAAATGAGATGTCACAGAGAGAGTGTGCGAAGTCCAGTCAAGACCGTG 1460
QY 1239 GAGAGAGACAG 1298
Db 1461 GAG 1496
QY 1299 GGAACAGAGAAATCTTGGCCCTGAGAAATCTGCTGAGCCCAAGAGAGTCCCCAGAA 1358
Db 1497 GAAACAGAGAGTGTGCGCAGAGAGAAATGTTGTGAATGATGACAACTCAGAGAA 1556
QY 1359 GCTGAGCCTGCTGAGAGCTGATGAAGAGAGAGAGATGTTGTCTGTGAGAGAGACAC 1418
Db 1557 GCCGAACCTGCAAGAGAGCTGTGATGACTCAAGAGAAAGTGTGTTCCGAGAGAGACCT 1616
QY 1419 ACTCACTGACAGACTTAAGTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1478
Db 1617 ACACAGAGAGCTGACCTCAAGCTGATGAAGAGTGTGTCTCAAAACCCCGAAGGGGTT 1676
QY 1479 GTCAGTGAAGTGAAGATGCTGTCTCTCAAGAAAGAAATCAAGGTAACAGGAGTCCCTG 1538
Db 1677 GTGAGTGAAGTGAAGATGCTGTCTCAAGAGAGAAATGAAGGTGAGAGAGAGTCCCTA 1736
QY 1539 AAGAACTCTTCACTAGCTCAAGGCTTAAAGAGGCTGTGGGAAAGAGAGAGAGAGAA 1598
Db 1737 AAGAACTTTTACAGACATGAGCTTAAAGAGTCTTGTGAAGAGAGAGAGAGAGAA 1796
QY 1599 CGAGAGGTGGGAG 1658
Db 1797 AGAGAG---GGAGAGAGAGAGAGATCAGGGAGAGACATCAGGTTCCAGCGGATTTCCG 1853
QY 1659 GAGAGTGTGATGAGAGAGAGAGAGAGAGTCTGCTGTCTCCCGAGAGAGCTGAGAG 1718
Db 1854 GACAGCAGAGAGAGAGAGAGAGAGAGAGAGAGTCTGCTCAATCCCTGAGAGAGCCGAGAG 1913
QY 1719 ACCAGTGTCTGAG 1778
Db 1914 ATCAAGTGTCTGAG 1973
QY 1779 ACTACTTCAGTGAAG 1838
Db 1974 GCTACTTCAGTGAAG 2033
QY 1839 ATGTGACACCAAG 1898
Db 2034 ATGTGAGAGCCCAAG 2093
QY 1899 GAGAGAGTCAAG 1958

Db 2094 GACAAGGTCAGAGGGCTACCTTGTCTTCCACGAGAGCAGCCTCTGAAATCCAAGAA 2153
Qy 1959 GAACTCAAAAACCTGTTGTGAGAAACAAAGCCAGAGAAACCAAGCGGTAGGGTGGATACT 2018
Db 2154 GAAATGAAGAGGAGCGTGAAGAGCCAAAGACCGGAAGAACCAAGCGCAAGGTGGATACC 2213
Qy 2019 TCAGTGTCTTTGGGAAGCACTGATTTTGTGTCGATCATCCAGAGAGAGCAGCAAGGCA 2078
Db 2214 TCAGTATCTTTGGGAAGCTTTAAATTTGTGTGGGATCATCCAAAGAAAGAGCAAGGAGAGG 2273
Qy 2079 TCCTCTTCAGATGATGAAGAGGGCCCAAGGACACTGGGAGGGGACAGTCAAGAGCAGAG 2138
Db 2274 TCCTCTTCAGTGAAGAGGGGACCAAAAGCAATGGAGAGACCACCAAGAAAGCTGAT 2333
Qy 2139 GAGGCCAGCAAGAACAAAGAACCCGGAACAGACGCTGTTCTTCAGACACCCAGAGCAG 2198
Db 2334 GAGGCCGGAAGAACAAAGAGACGGGACAGACGGGATCCTTGTGTTCCCAAGAACAT 2393
Qy 2199 GACCAAGGCAAGAAAGTTCTTCAACCCGAGCAGCAGGGGAAGCCCTTCGGAAGGGGAAGGT 2258
Db 2394 GATCCAGGGCAGGGAAGTTCTTCCCGAGCAAGCTGGAAGCCCTACCGAAGGGGGGGC 2453
Qy 2259 GTCTCCACTTTGGGAGTCATTTAAAAGATTAGTCACTCCAAAGAAAAAATCCAAGTCAAAA 2318
Db 2454 GTTTCACCTGGGAGTCATTTAAAAGTTAGTCAAGCCCAAGAAAAAATCAAGTCCAG 2513
Qy 2319 CTGGAAGAGAAAGCCGGAAGAC-----TCTAGTGTAGACGATGTTGCCACTGAG 2366
Db 2514 CTGGAAGAGAAAGCCGGAAGACTCCATAGCTGGGTCTGGTGTAGAACAATTCACCTCCAGAC 2573
Qy 2367 ATCGAAACGAGTAGAAGAAATCTTGGTTTCCATTTAAGAAATTCATCCCGGAGCGGG 2426
Db 2574 ACTGAACCCCGTAAAGAAAGAAATCTCGGTCTCAATCAAGAAAGTTTATTTCTGGACGAAG 2633
Qy 2427 AAGAAAAGGCGAGACGGGAAGAAACAAGCCACTGTGGAAGACTCAGGSCCAGTGGAG 2486
Db 2634 AAGAAAAGCCAGATGGGAACAAGAACAGCCCTGTTGAAAGACGAGGGGCCAACAGGG 2693
Qy 2487 ATAAATGAGGACACCTTAATGTCCCAGCGTGTGCTCTGTCTGATGATTAATGCAAGTG 2546
Db 2694 GCCAACCAAGATGACTCTGATGTCCCGCGGTGTCTCTGTCTGATGATGATGCTGTA 2753
Qy 2547 GAGGAGGAGAAAGTGG-----AAGCCAGGGGAATACGAGACTGCCAGCTGCTG 2597
Db 2754 GAAAGGAGAAATTTGAGGCGACAGCAAGCCCAAAAGGCGCAGAGCAGCCCGAGCAGAA 2813
Qy 2598 GGGCTGTGTACGTGTCCGAGGACTCAGTAAAGACTCTGGTCCACACTGTGAGTGTGCGCA 2657
Db 2814 GCAGCCACTGAGTGTCCAGGAGCTCAGCGAGAGTCAGGTTTCAATGATGGCAGCAGCT 2873
Qy 2658 GTCAATTGATGGACAGGCGAGTCACCAAGTGTGGAAGAGCGGTCTCCTTTCGTGATATCC 2717
Db 2874 GTCGCTGACGGGAGGCGAGCTACCATTTATGAAGAAAGGTCTCTCTTCTTGGATATCT 2933
Qy 2718 GCTTCCGTAAACAGAACTTTGAAACACAGAGCGGGAAGCCATGCCACCTGTTTGAAGAG 2777
Db 2934 GCTTCAGTGCAGAACTCTTTGAAACAAGTGAAGCTGAAAGCCGCACTGTTAACTGAGGAG 2993
Qy 2778 GTCACTCAAAAGACATCATTCG-----AGAGAAACTCTGTGCTCACCCAGACGTTACCA 2834
Db 2994 GTATTGGAAGAGAAAGTAAATTCAGAGAGAAAGAACCCCGGTTACTGAACTCTGCCA 3053
Qy 2835 GAGGGTAAAGATGCCCATGACGATGTCACAGTGAAGTGAAGTTTCACTCAGAAAGCT 2894
Db 3054 GAGAAACAGAGAGGCCCGGGGCGACACGGTCTGTAGTGAAGGGGAATTCACCCCGAAGCT 3113
Qy 2895 GTGACAGCCACAGAGACCTCAGAGGCTCTCCGTACTGAAGAAAGTTACCGAAGCATCCGGG 2954
Db 3114 GTGACAGCTGCAGAAAACCTGCAAGGGCCATTGGGTTCCGAAGAAAGGAAACCGAAGCATCTGCT 3173
Qy 2955 GCCGAAGAGACCAAGACATGTTGTCGCGAGTTTCCAGCTGACTGACTCCCGAGACACC 3014
Db 3174 GCTGAAGAGACCAAGAAAATGGTGTACAGCATGCTCCCGAGTTAACCCGATTCACCGACACCC 3233

Qy 3015 ACAGAGGAAGCCACCCAGTTTCAGGAGGTAGAGAGTGGTGTCTGTAGATACAGAAAGAG 3074
Db 3234 ACAGAGGAGGCCACTCCGCTGCAGAGGTTGGAAGGTGGCGTACTCTGCATAGAAGACAA 3293
Qy 3075 GAGGCCAGACGACGAGCCACTCTCCAGCCGTTTCGACAGAGGTGAAGAGGAGTCCCGAG 3134
Db 3294 GAGAGGCGGACTCAAGAGTCTCCAGGCTCTCCAGGCAAGTGAAGAAAGTGAAGAGGAATCCCG 3353
Qy 3135 GTCCCTGCACACC---CAGACTGTGCAGAGAACGGGGTCAAAAGCACTGGAGAAAGTTGAG 3191
Db 3354 CTGCTGGCACCGGTGGGCCAGAAAGATGTCTTCAGCCTGTGCAGAGAGCAGAGCCAGAA 3413
Qy 3192 GAGGTAGAGAGGACTCCGAAAGTGTGGCTTCGGAGAAAAGAGAGGACGTTATGCCGAAA 3251
Db 3414 AGACCAGAAAGCAGGCTGAAGCGTCCGGTCTGAAGAAAAGAGACGGATGTAGTGTGAAA 3473
Qy 3252 GGACCCGTGCAGGAACTGGAGCTGAGCATCTTCGACAGGGCTCTGAGACTGAGACGGCT 3311
Db 3474 GTAGATGCTCAGAGGGCAAAACTGAGCCTTTTACAAAGGGAAGGTGGTGGGCGACACC 3533
Qy 3312 ACTCCAGAGAGCCTTG---AAGTTCTTGAAGTCAAGGCGAGATGTAGAC-----CAT 3359
Db 3534 ACCCAGAAAGCTTTGAAAAGCTCTCAAGTCAAGAGAGCATGAGTCCAGTGAGCTT 3593
Qy 3360 GTCCGACAGTGCCA-----GGTTATCAAGCTCCAGCAGCTGATGGAA 3401
Db 3594 GTAACCACTGTCAAGCCGAAACCTTAGCTGGGGTAAATCAAGAGAGATGGTATGGAA 3653
Qy 3402 CAGGCCGTGGCCCTGAGTCACTCGAAACCTTCAGACAGTGAAGACAAATGGAGCACT 3461
Db 3654 CAGGCTATCCCGCTGACTCGGTGGAAACCCCTACAGACAGTGAAGTGTGTAAGACCC 3713
Qy 3462 CCCTTAGCAGATTCAAGACACTGCAGATGGGACACAGCAAGTGAACCATTTGACAGCCAG 3521
Db 3714 CCGTAGCCGACTTTGACGCAACAGGCAACAACCCAGAAACGAGATTTGGAATCCAT 3773
Qy 3522 GACAGTAAAGCCACTGCAGCTGTCCAGCTGCAGGCTCAAGGTCAAGAAAGAGGCGGCTACT 3581
Db 3774 GAGGAGAATCA---GGTGCATCTGGTACCAGTCAAGGCGCACAGAAAGCAGGCGAGTTCT 3830
Qy 3582 GCTCAGAAAGAGAGCCTTCGACACTACCTAATAATGTTCCAGCCCGAGGAAGAATCGG 3641
Db 3831 GCACAGAAAGAGAGGCTCCAGACCTTCCAGTTTGTGTTCCAGGAAGAACTAAAGAA 3890
Qy 3642 GAAGAACCAAGGAAGAGATGTTCTTGAACCTTACAGCAAGAGCTTACTGTCGACGCGTG 3701
Db 3891 CAATCAAGATGGAAGACACTCTAGAGCATACAGATAAGAGGTGTCAGTGGAAACTGTA 3950
Qy 3702 CCCGTTCTGGCAAGACTGAGGTGGGTCAAGAGGGTGA-----GGTTGACTGGTTG 3752
Db 3951 TCCATTCTGTCAAAGACTGAGGGGACTCAAGAGGCTGACCAAGTATGCTGATGAGAAACC 4010
Qy 3753 GATGGAGAAAAGTCAAGAAAGAACAGGAGGTGTTGTACACTCTGG-----ACCCAAC 3806
Db 4011 AAAGACGTACCAATTTTTCGAAGGACTTTGAGGGTCTATAGACACAGGCATAACAGTCA 4070
Qy 3807 AGTCAAAAGGCTGCTGATGTGACATATGACAGTGAAGTGTGGAGTGGCGGGTGTGAG 3866
Db 4071 CGGGAAGAGTCACTGAAGTTGCCCTTTAAAGGTGAAGGGGACAGAAAGTGAATGTAAA 4130
Qy 3867 GAAAAGGAGAGTACTGAAGTGCAGAG-----TCTTAGCCTGGAGGAGGGA 3911
Db 4131 AAGATGATGCTCTTGAACCTGCAGAGTCAAGCTAAGTCTCTCCATCCCCCGTGGAGAGA 4190
Qy 3912 GAGATGAAAACCTGACGTTTGAAGAGGAGAAAAGGAGACAAAGCCAGAGCAAGTGAAGTGA 3971
Db 4191 GAGATGTTAGTTCAAGTTCGAAGGAGGAGAAAACAGAAAGCAGAGCCCAACCCATGTGAATGA 4250
Qy 3972 GAAGG---TGACAGAGAAACAGCGCTCCTGAGCATGAAGGAACCTACGGGAAGCCAGTC 4028
Db 4251 GAGAAGCTTGAGCACGAAACAGCTGTTTACCGTATCTCTGAAGAGGTGCTAAGACGAGCTCTC 4310

Db	6387	TGTAAG--TTTTTGATTCTACTCTTATATGCTGGAGATCGCATTCACACATGGCATGAAATA	6444
Qy	6014	AGTCCTGCTCTTTTACAGATGGTATTTTGATAGATACTCGAGTGTCTGTGTTATATCTG	6073
Db	6445	AGTCAGGTCTCTTACAAATGGTATTTTGATAGATACTGGATTGTGCTTGTGCCATATTG	6504
Qy	6074	TGCCCCCTCTTTTAAGAACAAATGTCGATTATGTTCTCTTGGATAAAATTTGCTATTGCAAA	6133
Db	6505	TGCCATTTCCCTTTAAGAACAAATGTTGCAACACACATTCATTGGATAAGTTGTGATTGACGA	6564
Qy	6134	CTGATTTAAATAAA	6147
Db	6565	CTGATTTAAATAAA	6578

```

RESULT 4
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463

```

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE: 26-AUG-1991
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: ptz5pt-F18
US-08-232-463-14

```

Query Match      1.6%, Score 99.8, DB 1, Length 7218;
Best Local Similarity 2.1%, Pred. No. 1.5e-15;
Matches 8; Conservative 262; Mismatches 109; Indels 0; Gaps 0;

QY 944 AAAAGAGATGATCTGGAAATCGCGAAGAGAAAGGACGAGGAGAAAAGTAGA 1003
      ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 1431 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1372
      ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
QY 1004 CGAGGAGAAAAGAAAAGACAGAGCCAGCTCGGAGGAGGAGGAGCCGCGAAGACAC 1063

```

```

Db      1371 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : ~~~~~~
Qy     1064 AGACAGCCGAGTTGTCA GACA GACTAG CAGA ACGTGG AGCTGC CTTTGA GAAC CACG GT   1123
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : ~~~~~~
Db     1311 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR : : : : : : : : : : : : : : : : : : : : ~~~~~~
Qy    1124 TGGTGACCTGGAGCATCGTCAGAGGAGAACTGTGCTCCTTTGCCAACGGAAGTCTTTGA   1183
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : ~~~~~~
Db     1251 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR : : : : : : : : : : : : : : : : : : : : ~~~~~~
Qy    1184 TGAGAAGATGGAACCCACC AACGAAG TTGTTG CAGAGG TCACGTG ACAGCAC CGTGGAG AA   1243
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : ~~~~~~
Db     1191 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR : : : : : : : : : : : : : : : : : : : : ~~~~~~
Qy    1244 GACAGAGGAGCAGGAGGAGGAGGAGGAGGAGCTGAAGGGGGCGCTGGTGGTAGAAGGAAC   1303
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : ~~~~~~
Db     1131 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR : : : : : : : : : : : : : : : : : : : : ~~~~~~
Qy    1304 AGGAGAATCTTGCCCCCT   1322
       : : : : | . ||| |||| 
Db     1071 RRRRRATCGCAAGTCCCT   1053

RESULT 5
US-09-249-585A-4/c
; Sequence 4, Application US/09249585A
; Patent No. 6417002
; GENERAL INFORMATION:
; APPLICANT: Horlick, Robert
; TITLE OF INVENTION: METHOD
; FILE REFERENCE: 0867/OD905
; CURRENT APPLICATION NUMBER: US/09/249,585A
; CURRENT FILING DATE: 1999-02-11
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 4
; LENGTH: 1926
; TYPE: DNA
; ORGANISM: Epstein Barr Virus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1926)
; OTHER INFORMATION: template strand of EBNA-1 DNA
US-09-249-585A-4
```

Query Match	1.2%	Score	71.2;	DB	4;	Length	1926;		
Best Local Similarity	47.7%	Pred.	No. 1.9e-08;						
Matches	208;	Conservative	0;	Mismatches	228;	Indels	0;	Gaps	0;
Qy	980	GGAGCAAGAGCAGAAAAGTAGACGAGGAGAAAAGGAAAAGACAGAGCCAGCCTCGGA	1039						
Db	977	GGAGGACGAGGACCGGGAGGACGAGGACCGGGAGGAGACGAGGACGGGGA	918						
Qy	1040	GGAGCAGGAGCCGCGAGAAGACACAGACCAGGCCAGGTTGTTCAGCAGACTACGAGAAGGT	1099						
Db	917	GGACGAGACCGGGAGGACCGGGAGGACGAGGACCGGGAGGACGAGGACGG	858						
Qy	1100	GGAGTGCTCTTTGGAAGACCAAGTTGTGTGACCTCGAGGCATCTGCAGAGGAGAAAGTGTGC	1159						
Db	857	CGAGGACCGGGAGGAGGACGAGGACCGGGAGGACCGGGAGGAGGACGAGGACCGGGAGGA	798						
Qy	1160	TCCTTTGCAACGGAAGTGTGTGATGAGAATCGGAGCCACCAGAAAGTTGTCAGA	1219						
Db	797	CGGGAGGACGAGGACCGGGAGGACCGGGAGGACGAGGACCGGGAGGACCGGGAGGACGA	738						
Qy	1220	GCTCCACTGTAGCACCTGTGGAAGAAGACAGAGGAGGAGCAGGAGGAGGAGGAGGCTGA	1279						
Db	737	GGAACGGGAGGACGAGGACCGGGAGGACGAGGACCGGGAGGACGAGGACCGGGAGGACGG	678						
Qy	1280	AGGGGGCGTGTGTAGAAGGAACAGGAGAACTCTTGTCCCTCGAGAACTGGCTGAGCC	1339						
Db	677	CGAGGACCGGGAGGAGGACGAGGACCGGGAGGACGAGGACCGGGAGGACCGGGAGGACGG	618						


```

RESULT 8
US-09-050-863-2
; Sequence 2, Application US/09050863
; Patent No. 6114111
; GENERAL INFORMATION:
; APPLICANT: Lao, Ying
; APPLICANT: Hiang, Betty
; APPLICANT: Pavan, Don
; TITLE OF INVENTION: Mammalian Protein Interaction Cloning
; TITLE OF INVENTION: System
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/050,863
; FILING DATE: 30-MAR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-65638/DJB/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 949-8711
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2580 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA
US-09-050-863-2

Query Match          1.1%; Score 70.8; DB 3; Length 2580;
Best Local Similarity 45.3%; Pred.No.2.8e-08;
Matches 335; Conservative 0; Mismatches 402; Indels 3; Gaps 2;

QY   946  AAGAGGATGATCTGGAAACTGCCAGAGAAGAAAAGGACGAAGAGCGACAAAAAAGTAGACG 1005
      |||
DB    716  AGGAGAGGGGCAGGAGGGGCAGGAGGGGCAGGAGGGGCAGGAGGAGGAGGGGCAGG 775
      |||
QY   1006 AGGAAGAAAAGGAAAAGAACAGCAGCCGTCGGAGGAGCAGGAGCCGCGACGAAGACACAG 1065
      |||
DB    776  AGCAGGAGGAGGGGCAGGAGGGGCAGGAGGGGCAGGAGGAGGAGGGGCAGGAGGAGG 835
      |||
QY   1066 ACCAGGCCAGGTGTGCACGACACTACGAGAGAGTGGAGCTGCCCTTTGGAACACAGGTTG 1125
      |||
DB    836  AGGAGGGCAGAGGGGCAGGAGGAGGAGGGGCAGGAGGGGCAGGAGGGGCAGGAGGAGG 895
      |||
QY   1126 GTGACCTGGAGGCATCGTCAGAGGAGAGTGCTGTCTCTTTGGCAACGGAAGTGTTCATG 1185
      |||
DB    896  AGGAGGGGGGCAGGAGCAGGAGGAGGGGCAGGAGGGGCAGGAGGAGGAGGGGCAGG 955
      |||
QY   1186 AGAAGATGGAAGCCCAACAAGAAGTTGTCAGAGGTCACGTGAGCACCGTGAGGAAGA 1245
      |||
DB    956  AGGGGCAGGAGGGGCAGGAGCAGGAGGAGGGGCAGGAGGAGGAGGGGCAGGAGGAGGAGG 1015
      |||
QY   1246 CAGAGGAGGAGCAGGGAGGAGGAGGAGGCTGAAGGGGGCGTGTGCTAGAACGAACAG 1305
      |||
DB    1016 AGGGCGAGGAGCA--GGAGGGCAGAGCAGGAGGAGGGGCAGGAGGGGCAGGAGGGGCAG 1074
      |||
QY   1306 GAGAACTCTTGC CCCCTTGAGAAACTGGCTGAGCCCCCAGGAGGCTCCCCCAGGAAGCTGAGC 1365

```

US-09-359-081-2

Query Match 1.1%, Score 70.8; DB 4; Length 2580;
Best Local Similarity 45.3%; Pred. No. 2.8e-08;
Matches 335; Conservative 0; Mismatches 402; Indels 3; Gaps 2;

QY 946 AAGAGATGATCTGGAAACTGCGGAGAGAAAGAGACCAAGAGCAAAAAGTGAACG 1005
DB 716 AGGAGAGGGGCGAGAGGGGCGAGGGGCGAGAGGCGGCGAGAGGAGGAGGGCGAGG 775
QY 1006 AGGAG 1065
DB 776 AGGAG 835
QY 1066 ACCAGGAG 1125
DB 836 AGGAG 895
QY 1126 GTGACCTGAG 1185
DB 896 AGGAG 955
QY 1186 AGGAG 1245
DB 956 AGGAG 1015
QY 1246 CAG 1305
DB 1016 AGGAG 1074
QY 1306 GAGAACTCTGCCCCCTGAGAACTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1365
DB 1075 GAGCA--GAG 1132
QY 1366 CTGCTGAG 1425
DB 1133 AGGAG 1192
QY 1426 TGACAG 1485
DB 1193 AGCAG 1252
QY 1486 AGCTGAG 1545
DB 1253 AGGAG 1312
QY 1546 TCTTCAGTAGCTCAGAGCTTAAAGAGCTGCTGGAGAGAGAGAGAGAGAGAGAGAG 1605
DB 1313 GCGAG 1372
QY 1606 GTGGGAG 1665
DB 1373 TCGAG 1432
QY 1666 CTGATGAG 1685
DB 1433 TCGAG 1452

RESULT 10
US-09-130-114-1/c
; Sequence 1, Application US/09130114
; Patent No. 5976807
; GENERAL INFORMATION:
; APPLICANT: Horlick, Robert A.
; APPLICANT: Damaj, Bassem B.
; APPLICANT: Robbins, Alan K.
; TITLE OF INVENTION: Eukaryotic Cells Stably Expressing Genes
; FILE REFERENCE: 0867/ID903US1
; CURRENT APPLICATION NUMBER: US/09/130,114
; CURRENT FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 36

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 1
; LENGTH: 5452
; TYPE: DNA
; ORGANISM: YEENA
US-09-130-114-1

Query Match 1.1%, Score 70.8; DB 2; Length 5452;
Best Local Similarity 45.3%; Pred. No. 4.3e-08;
Matches 335; Conservative 0; Mismatches 402; Indels 3; Gaps 2;

QY 946 AAGAGATGATCTGGAAACTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1005
DB 2089 AGGAG 2030
QY 1006 AGGAG 1065
DB 2029 AGCAG 1970
QY 1066 ACCAGGAG 1125
DB 1969 AGGAG 1910
QY 1126 GTGACCTGAG 1185
DB 1909 AGGAG 1850
QY 1186 AGGAG 1245
DB 1849 AGGAG 1790
QY 1246 CAG 1305
DB 1789 AGGAG 1731
QY 1306 GAGAACTCTGCCCCCTGAGAACTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1365
DB 1730 GAGCA--GAG 1673
QY 1366 CTGCTGAG 1425
DB 1672 AGGAG 1613
QY 1426 TGACAG 1485
DB 1612 AGCAG 1553
QY 1486 AGCTGAG 1545
DB 1552 AGGAG 1493
QY 1546 TCTTCAGTAGCTCAGAGCTTAAAGAGCTGCTGGAGAGAGAGAGAGAGAGAGAGAG 1605
DB 1492 GCGAG 1433
QY 1606 GTGGGAG 1665
DB 1432 TCGAG 1373
QY 1666 CTGATGAG 1685
DB 1372 TCGAG 1353

RESULT 11
US-08-910-647-1
; Sequence 1, Application US/08910647
; Patent No. 6251433
; GENERAL INFORMATION:
; APPLICANT: Zuckermann et al.
; TITLE OF INVENTION: Compositions and Methods for
; Polynucleotide Delivery
; NUMBER OF SEQUENCES: 4

[illegible]

QY 1066 ACCAGCCAGGTTGTCACAGACTACGAGAAAGTGGAGCTGCTTTGGAGACAGGTTG 1125
 Db 882 AGGAGGGGCGAGAGGGGCGAGAGCGAGAGGGGCGAGAGGGGCGAGAGGCG 941
 QY 1126 GTGACCTGAGGCACTGTCAGAGAGAGTGTCTCTTTGGCAACGAAAGTGTGATG 1185
 Db 942 AGGAGAGGGGCGAGAGCGAGAGAGGGGCGAGAGGGGCGAGAGCGAGAGGGGCG 1001
 QY 1186 AGAAGTGAAGCCCAAGAAAGTTGTCAGAGGTCACCTGAGCACCTGTGAGAAAG 1245
 Db 1002 AGGGGCGAGAGGGGCGAGAGCGAGAGGGGCGAGAGCGAGAGGGGCGAGAGGGGCG 1061
 QY 1246 CAGAGAGAGCGAGAGAGAGAGAGGCTGAAGGGGGCGTGGTGAAGAAAGAACG 1305
 Db 1062 AGGGGCGAGAGCA-GGAGGGGCGAGAGCGAGAGGGGCGAGAGGGGCGAGAGGGGCG 1120
 QY 1306 GAGAACTCTTGGCCCCCTGAGAACTGCTGAGCCCCAGAGAGTCTCCCGAGAAAGCTGAGC 1365
 Db 1121 GAGCA--GAGAGGGGCGAGAGCGAGAGGGGCGAGAGCGAGAGGGGCGAGAGGGGCG 1178
 QY 1366 CTGCTGAGAGAGCTGATGAAGAGAGAGAGTGTCTCTTGAAGAGACCACTCAAC 1425
 Db 1179 AGGAGGGGCGAGAGCGAGAGGGGCGAGAGGGGCGAGAGCGAGAGGGGCGAGAGGGGCG 1238
 QY 1426 TGACAGACTTAAGTCTCTGAAGAGAAAGCCCTGCCAAGACCCAGAGGCAATTTCAATG 1485
 Db 1239 AGCAGAGAGAGGGGCGAGAGGGGCGAGAGCGAGAGGGGCGAGAGGGGCGAGAGGGGCG 1298
 QY 1486 AGGTGAGAGTGTCTCTCTCAAGAAAGATCAAGTACAGGAAAGTCTTGAAGAAAC 1545
 Db 1299 AGGGGCGAGAGGGGCGAGAGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCG 1358
 QY 1546 TCTTCAGTAGCTCAGGCTTAAAGAACTGTCTGGGAGAGCAAGAGGGGAAACAGAGAG 1605
 Db 1359 GGCAGAGAGAGAGAGGGGCGAGAGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCG 1418
 QY 1606 GTGGGGGAGACGAAGAGCCTTGAAGATACCAACACTTCAACCCGAATCCCAAGAGAGTG 1665
 Db 1419 TCAGAGAGCACTGAGAGCGGGGCTCAGAGAGTAGTGAAGCGGGGCTCAGAGAGGTAG 1478
 QY 1666 CTGATGACGAGAGAGAGAG 1685
 Db 1479 TGAAGCGCGCGGGGTAGAG 1498

RESULT 13

US-07-884-811-15
 ; Sequence 15, Application US/07884811

; GENERAL INFORMATION:

; APPLICANT: Godowski, Paul J. Lokker, Nathalie A. Mark, Melanie R.
 ; TITLE OF INVENTION: SINGLE-CHAIN HEPATOCYTE GROWTH FACTOR VARIANTS
 ; NUMBER OF SEQUENCES: 21
 ; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Genentech, Inc.
 ; STREET: 460 Point San Bruno Blvd
 ; CITY: South San Francisco
 ; STATE: California
 ; COUNTRY: USA

; ZIP: 94080

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: patin (Genentech)

; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/884,811

; FILING DATE: 19920518

; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:

NAME: Dreyer, Ginger R.
 REGISTRATION NUMBER: 33,055
 REFERENCE/DOCKET NUMBER: 755.1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415/225-3216
 TELEFAX: 415/952-9881
 TELEX: 910/371-7168
 INFORMATION FOR SEQ ID NO: 15:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 10596 bases
 TYPE: NUCLEIC ACID
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-07-884-811-15

Query Match 1.1%; Score 70.8; DB 1; Length 10596;
 Best Local Similarity 45.3%; Pred. No. 6.3e-08;
 Matches 335; Conservative 0; Mismatches 402; Indels 3; Gaps 2;

QY 946 AAGAGATGATCTGTGAAGAACTGCCGAGAAAGAGCAAGAGCGAAGAAAGTACAGC 1005
 Db 2296 AGGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCG 2355
 QY 1006 AGAAGAAAGAAAGAAAGCAAGAGCCAGCTCTGAGAGCGAGAGCCGCGAAGAGACAG 1065
 Db 2356 AGCAGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGCGAGAGAGAGGGGCGAGAGGGGCG 2415
 QY 1066 ACCAGCCAGGTTGTACAGACACTACGAAAGTGTGAGCTCTTTGAAAGACAGGTTG 1125
 Db 2416 AGGAGGGGCGAGAGGGGCGAGAGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCG 2475
 QY 1126 GTGACCTGAGAGCATCTGTCAGAGAGAGTGTCTCTTTGGCAACGAAAGTGTGATG 1185
 Db 2476 AGGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAG 2535
 QY 1186 AGAAGTGAAGCCCAAGAGAGTGTGTCAGAGGTCACGTGAGCACTTGAAGAAAG 1245
 Db 2536 AGGGGCGAGAGGGGCGAGAGCGAGAGGGGCGAGAGCGAGAGGGGCGAGAGGGGCGAGAGGGGCG 2595
 QY 1246 CAGAGAGAGAGAGAGAGAGAGAGAGCTGAAGGGGGCGTGTGTAGAGAGAGAG 1305
 Db 2596 AGGGGCGAGAGCA-GAAGGGGCGAGAGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAG 2654
 QY 1306 GAGAACTCTTCCCCCTGAGAACTGGCTGAGCCCCAGAGAGTCCCCAGAGAGCTGAGC 1365
 Db 2655 GAGCA--GAGAGGGGCGAGAGCGAGAGGGGCGAGAGCGAGAGGGGCGAGAGGGGCG 2712
 QY 1366 CTGCTGAGAGACTGATGAAGAGAGAGATGTGTCTCTGAGAGAGACCACTCAAC 1425
 Db 2713 AGGAGGGGCGAGAGCGAGAGGGGCGAGAGGGGCGAGAGCGAGAGGGGCGAGAGGGGCGAG 2772
 QY 1426 TGACAGACTTAAGTCTTGAAGAAAGACGCTGCCAAGACCCAGAGAGCAATTTCACTG 1485
 Db 2773 AGCAGAGAGAGGGGCGAGAGGGGCGAGAGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAG 2832
 QY 1486 AGGTGAGAGTGTCTCTCTCAGGAAAGATCAAGGTACAGGAAAGTCTTGAAGAAAC 1545
 Db 2833 AGGGGCGAGAGGGGCGAGAGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAG 2892
 QY 1546 TCTTCAGTAGCTCAGGCTTAAAGAACTGTCTGGGAGAGCGAGAGGGGAAACAGAGAG 1605
 Db 2893 GGCAGAGAGCGAGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGCGAGAGGGGCGAGAGGGGCG 2952
 QY 1606 GTGGGGGAGACGAAGAGCCTTGAAGATACCAACATTTCAACCGAATCCCCAGAGAGTG 1665
 Db 2953 TCAGAGAGCACTGAGAGCGGGGCTCAGAGAGTAGTGAAGCGGGGCTCAGAGAGGTAG 3012
 QY 1666 CTGATGACGAGAGAGAG 1685
 Db 3013 TGAAGCGCGCGGGGTAGAG 3032

RESULT 14

```

US-07-885-971-15
; Sequence 15, Application US/07885971
; Patent No. 5328837
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J. Lokker, Nathalie A. Mark, Melanie R.
; TITLE OF INVENTION: HEPATOCYTE GROWTH FACTOR PROTEASE DOMAIN VARIANTS
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/885,971
; FILING DATE: 19920518
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dregler, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: 779
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-3216
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10596 bases
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-07-885-971-15

Query Match 1.1%; Score 70.8; DB 1; Length 10596;
Best Local Similarity 45.3%; Pred. No. 6.3e-08;
Matches 335; Conservative 0; Mismatches 402; Indels 3; Gaps 2;

Qy 946 AAGAGGATGATCTGAAACTGCCGAGAGAGAAAAGGACAGAGCCAGAAAAGTAGACG 1005
Db 2296 AGGAGGGGGCGAGGAGGGCGAGGAGGGCGAGGAGGGCGAGGAGGGGGCAGG 2355
Qy 1006 AGGAAGAAAAGGAAAAGACAGAGCCAGCTCGAGGAGCAGAGCCGGCAGAAAGACACAG 1065
Db 2356 AGCAGGAGGGGGCGAGGAGGGCGAGAGGGCGAGAGGAGGAGGGGGCAGGAGCAGG 2415
Qy 1066 ACCAGGCCAGGTTGTACAGCAGACTACGAGAAGGTGGAGCTGCTTTGGAAGACCAGGTTG 1125
Db 2416 AGGAGGGGCGAGGAGGGCGAGGAGCAGGAGGAGGGCGAGGAGGGGGCAGGAGC 2475
Qy 1126 GTGACCTGGAGGCATCGTCAGAGGAGAAGTGTGCTCTTTTGCAACGAGNAGTGTGTATG 1185
Db 2476 AGGAGGGGGCGAGGAGCAGGAGGAGGGGGCGAGGAGGGGGCGAGGAGGGGGCAGG 2535
Qy 1186 AGAAGTGGAAAGCCCAACAAAGATTGTTGCAGAGGTCCACGTGAGCACCGTGGAGAGA 1245
Db 2536 AGGGCGAGAGGGGCGAGAGCAGGAGGAGGGCGAGGAGGGGGCGAGGAGGGGGCAGG 2595
Qy 1246 CAGAGGAGGAGCAGGAGGAGGAGGAGGCTGAAGAGGGGGCGTGTGTGTAGAAAGGAA 1305
Db 2596 AGGGCGAGGAGCA-GGAGGGGCGAGGAGCAGGAGGAGGAGGGGGCGAGGAGGGGGCAG 2654
Qy 1306 GAGAATCTTGCCCTCTGAGAAACTGGGTGAGCCCCAGAGAGTCCCCCAGGAAGCTGAGC 1365
Db 2655 GAGCA--GGAGGGGCGAGGAGCAGGAGGGGGCGAGGAGCAGGAGGGGGCGAGGAGGGG 2712

```

Best Local Similarity 45.3%; Pred. No. 6.3e-08;
Matches 335; Conservative 0; Mismatches 402; Indels 3; Gaps 2;

```

Qy 946 AAGAGATGATCTGTGAAATCTCGAAGAAAGAAAGACCAAGGACAGAAAAAGTAGACG 1005
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2296 AGGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGG 2355
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1006 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1065
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2356 AGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2415
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1066 ACCAGGCGAGGTTGTCTACAGACTACGAGAAAGTGTGAGCTGCTTGGAAAGCAGATTG 1125
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2416 AGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2475
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1126 GTGACCTGTGAGGCAATCTCAGAGGAGAAAGTGTCTCTTTGGCAACGAAAGTTGATG 1185
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2476 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2535
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1186 AGAAGATGAAAGCCCAAGAAAGTTTGTTCAGAGGTCACCTGTAGCACCTGTGAAAGA 1245
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2536 AGGGGCAAGAGGGGCGAGGAGCGAGGAGGGGCGAGGAGCGAGGAGGGGCGAGG 2595
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1246 CAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1305
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2596 AGGGGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2654
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1306 GAGAACTCTTGCCCTGAGAAATGAGCTGAGCCCGAGAGGTTCCCGAGAAAGCTGAGC 1365
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2655 GAGCA--GGAGGGGCGAGGAGCGAGGAGGGGCGAGGAGCGAGGAGGGGCGAGG 2712
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1366 CTGCTGAGGAGCTGATGAAAGCAGAGAGATGTGTCTCTTGAGAGACCACTCAAC 1425
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2713 AGGAGGGCGAGGAGCGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGG 2772
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1426 TGACAGACTTAAGTCTTGAAGGAAGCGCTGCCCAACCCAGAGAGGCAATTGTCAGTG 1485
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2773 AGCAGAGAGAGGGGCGAGGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2832
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1486 AGGTGAGATGCTGTCTCTCAGAAAGATCAAGTACAGGAAATCCCTTGAAGAAAC 1545
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2833 AGGGGCGAGGAGGGGCGAGGAGCGAGGAGGGGCGAGGAGCGAGGAGGGGCGAGG 2892
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1546 TCTTCAGTAGCTCAGGCTTAAAGAACTGTCTTGGAAAGACAGAAAGGGAACAGAGG 1605
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2893 GGCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2952
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1606 GTGGGGGAGACGAAAGCCTTGAGAAATCCAAACATTCAACCCGAATCCCAAGAGTG 1665
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2953 TCGAGGAGGCACTGAGAGGCGGGGCTCGAGAGGTAGTGAGGCGGGGTCGAGAGGTTAG 3012
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1666 CTGATGAGCAGAAAGGAGAG 1685
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3013 TGGAGGCCCGCGGGGTAGAG 3032
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Search completed: December 12, 2002, 14:19:35
Job time : 351 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 12, 2002, 16:19:50 ; Search time 244 Seconds
(without alignments)
9983.770 Million cell updates/sec

Title: US-09-902-432-3
Perfect score: 6160
Sequence: 1 cgggcccggggagtagaag.....aaataaacatatttgactac 6160

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 355320 seqs, 197730502 residues

Total number of hits satisfying chosen parameters: 710640

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6160	100.0	6160	9 US-09-902-432-3	Sequence 3, Appli
2	4831.4	78.4	5134	9 US-09-902-432-1	Sequence 1, Appli
3	2144.4	34.8	6608	10 US-09-738-877-1	Sequence 1, Appli
4	2144.4	34.8	6608	10 US-09-880-107-3439	Sequence 3439, Ap
5	1945.6	31.5	5345	10 US-09-738-877-2	Sequence 2, Appli
6	550	8.9	563	10 US-09-917-800A-1223	Sequence 1223, Ap
7	162.2	2.6	178	10 US-09-917-800A-1475	Sequence 1475, Ap
8	106	1.7	720	9 US-10-101-487-74	Sequence 74, Appl
9	106	1.7	720	9 US-10-101-487-76	Sequence 76, Appl
10	89.6	1.5	522	9 US-10-101-487-71	Sequence 71, Appl
11	89.6	1.5	530	9 US-10-101-487-73	Sequence 73, Appl
12	89.6	1.5	554	9 US-10-101-487-69	Sequence 69, Appl
13	89.6	1.5	554	9 US-10-101-487-106	Sequence 106, App
14	71	1.2	575	10 US-09-864-761-20733	Sequence 20733, A
15	71	1.2	1969	10 US-09-864-761-3972	Sequence 3972, Ap
16	65.8	1.1	659158	9 US-09-771-208-20	Sequence 20, Appl
17	64.4	1.0	100	10 US-09-728-445-335	Sequence 336, App
18	61	1.0	700	10 US-09-864-761-17529	Sequence 17529, A
19	59.4	1.0	53226	10 US-09-818-264-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1

US-09-902-432-3
; Sequence 3, Application US/09902432
; Patent No. US20020160002A1
; GENERAL INFORMATION:
; APPLICANT: Irwin H. Gelman
; APPLICANT: Susan G. Jaken
; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
; FILE REFERENCE: A30558-A-FWC-A 070156.0597
; CURRENT APPLICATION NUMBER: US/09/902,432
; CURRENT FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 08/978,277
; PRIOR FILING DATE: 1997-11-25
; PRIOR APPLICATION NUMBER: 08/665,401
; PRIOR FILING DATE: 1996-06-18
; PRIOR APPLICATION NUMBER: 08/635,121
; PRIOR FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 6160
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; US-09-902-432-3

Query Match	100.0%	Score 6160;	DB 9;	Length 6160;
Best Local Similarity	100.0%	Pred No. 0;		
Matches 6160;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	CGGGCCGGGGAGTAGAAGAGCCACTGAGCCATGGCGGAGGAGTGGTCTAGTGGCCATG	120	
Db	1	CGGGCCGGGGAGTAGAAGAGCCACTGAGCCATGGCGGAGGAGTGGTCTAGTGGCCATG	60	
Qy	61	GGAGCCCGGAGCAGCCGGGGGAGCGACACCGCGAGGAGTGGTCTAGTGGCCATG	120	
Db	61	GGAGCCCGGAGCAGCCGGGGGAGCGACACCGCGAGGAGTGGTCTAGTGGCCATG	120	
Qy	121	GGCCCGGAGTGAAGCCCTCGGGAGCAGCTGGAGACCCCGCGAGCCGCCACCA	180	
Db	121	GGCCCGGAGTGAAGCCCTCGGGAGCAGCTGGAGACCCCGCGAGCCGCCACCA	180	
Qy	181	AGTCTCCACAGAAAGATGGCCAGCTGTTTCTGTCAACGGCGTAGCTAGTGAACAGGAGATG	240	

c	20	59.2	1.0	511	10	US-09-864-761-22817	Sequence 22817, A
c	21	57.2	0.9	475	10	US-09-864-761-6203	Sequence 6203, Ap
c	22	56.2	0.9	474	10	US-09-864-761-3470	Sequence 3470, Ap
c	23	55.4	0.9	2142	10	US-09-917-800A-1567	Sequence 1567, Ap
c	24	54.2	0.9	334	10	US-09-864-761-20240	Sequence 20240, A
c	25	53.8	0.9	267	10	US-09-864-761-27984	Sequence 27984, A
c	26	53.8	0.9	474	10	US-09-864-761-11284	Sequence 11284, A
c	27	53.8	0.9	3809	12	US-10-001-870-68	Sequence 68, Appl
c	28	52.8	0.9	726	10	US-09-864-761-19707	Sequence 19707, A
c	29	52.8	0.9	1951	10	US-09-864-761-2926	Sequence 2926, Ap
c	30	52.4	0.9	2623	9	US-09-835-233-3	Sequence 3, Appli
c	31	50.6	0.8	623	9	US-09-804-409A-7	Sequence 7, Appli
c	32	50.4	0.8	5387	9	US-10-001-873-22	Sequence 22, Appl
c	33	50.2	0.8	438	10	US-09-864-761-4988	Sequence 4988, Ap
c	34	49.6	0.8	663	9	US-10-001-857-11	Sequence 11, Appl
c	35	49.4	0.8	659158	9	US-09-771-208-20	Sequence 20, Appl
c	36	49	0.8	1728	9	US-09-906-514-3	Sequence 3, Appli
c	37	48.8	0.8	360	10	US-09-864-761-28274	Sequence 28274, A
c	38	48.8	0.8	450	10	US-09-864-761-11696	Sequence 11696, A
c	39	48.4	0.8	700	9	US-10-040-739-1385	Sequence 1385, Ap
c	40	48.4	0.8	768	9	US-09-938-842A-812	Sequence 812, App
c	41	48.4	0.8	4316	10	US-09-880-107-3713	Sequence 3713, Ap
c	42	48	0.8	276	10	US-09-864-761-20595	Sequence 20595, A
c	43	48	0.8	342	10	US-09-822-263-11	Sequence 11, Appl
c	44	48	0.8	555	10	US-09-822-263-3	Sequence 3, Appli
c	45	48	0.8	742	10	US-09-770-149-52	Sequence 52, Appl

181 AGCTCCACAGAAAGATGGCCAGCTGTCTTGTTCACGGCTAGCTGAACAAGAGATG 240
241 TCCATGTCCAGAGAGAAACCAGAGGGGCAAGAGAAAGTCGTTGATGAGATGTTG 300
242 TCCATGTCCAGAGAGAAACCAGAGGGGCAAGAGAAAGTCGTTGATGAGATGTTG 300
301 GACAGCGAGATCAGAAAGTGTGAGAGAAAAAGACCGAGTTGAGAAATGSGCGCACT 360
302 GACAGCGAGATCAGAAAGTGTGAGAGAAAAAGACCGAGTTGAGAAATGSGCGCACT 360
361 CCAGAGCTGTGAGAGATTCACAAAGATGGGAGAGAGACATCAAGAAATTAATTGAC 420
362 CCAGAGCTGTGAGAGATTCACAAAGATGGGAGAGAGACATCAAGAAATTAATTGAC 420
421 AGATCCCTGCTTCAGAAAAACAATGTGAGAAATGTGAACAGCTGCTGATGCCAGCTA 480
422 AGATCCCTGCTTCAGAAAAACAATGTGAGAAATGTGAACAGCTGCTGATGCCAGCTA 480
481 ATGATGTGGCTTCAGAAAGTATTAAATTTGTTTAAATTCACGGTGAAGAG 540
482 ATGATGTGGCTTCAGAAAGTATTAAATTTGTTTAAATTCACGGTGAAGAG 540
541 ATAAAAATGAAAGTCAATCTGTCCAACTACTCACTGCAAGAAAGATGAAGGCAAG 600
542 ATAAAAATGAAAGTCAATCTGTCCAACTACTCACTGCAAGAAAGATGAAGGCAAG 600
601 GGGAGAAAGCTCTGTCCGAGCTGAGACCAACAGAGCCCAAGTGTGAGACTGCGCTG 660
602 GGGAGAAAGCTCTGTCCGAGCTGAGACCAACAGAGCCCAAGTGTGAGACTGCGCTG 660
661 GAGAGTCAGATCCAAAGAAAGTGAAGCAATCCACAGAGAGCAAGAGCAAGCAAGC 720
662 GAGAGTCAGATCCAAAGAAAGTGAAGCAATCCACAGAGAGCAAGAGCAAGCAAGC 720
721 TGAAGCAAGAACAGAGCAAGCAAGAAATCCCTTCAGAGCCGAATCTGATCAAGGGCTG 780
722 TGAAGCAAGAACAGAGCAAGCAAGAAATCCCTTCAGAGCCGAATCTGATCAAGGGCTG 780
781 AGAAGAAAGCCAAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
782 AGAAGAAAGCCAAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
841 AATCCCGAGAGCCAGTCAACAGTGAACAACATCTTCTTCAAGAGTTCTTCACTC 900
842 AATCCCGAGAGCCAGTCAACAGTGAACAACATCTTCTTCAAGAGTTCTTCACTC 900
901 ACGGTGGGGCGGCTGGGGCAAGAGACAGCTTCAAGAAATCAAAAGAGAGTGTCTGG 960
902 ACGGTGGGGCGGCTGGGGCAAGAGACAGCTTCAAGAAATCAAAAGAGAGTGTCTGG 960
903 ACGGTGGGGCGGCTGGGGCAAGAGACAGCTTCAAGAAATCAAAAGAGAGTGTCTGG 960
961 AAACTGCGGAG 1020
962 AAACTGCGGAG 1020
1021 AGACAGAGCCAGCTCGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
1022 AGACAGAGCCAGCTCGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
1081 CAGCAGATTAAGAGAGAGTGAAGTGTCTTGGAGAGACAGAGTGTGAACCTGGAGGCA 1140
1082 CAGCAGATTAAGAGAGAGTGAAGTGTCTTGGAGAGACAGAGTGTGAACCTGGAGGCA 1140
1141 CGTCAGAGAGAGAGTGTCTCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
1142 CGTCAGAGAGAGAGTGTCTCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
1201 ACCAAGAGAGTGTGTGAGAGAGTCAAGTGAAGCAAGTGAAGAGAGAGAGAGAG 1260
1202 ACCAAGAGAGTGTGTGAGAGAGTCAAGTGAAGCAAGTGAAGAGAGAGAGAGAG 1260
1261 GAGAGAGAGAGAGAGTGAAGGGGCGTGTGTGAAGAGAGAGAGAGAGAGAGAGAG 1320

1261 GAGAGAGAGAGAGAGCTGAAGGGGGCGTGTGTGAAGAGAGAGAGAGAGAGAGAG 1320
1321 CTGAGAAACTGGCTGAGCCCCAGAGAGTCCCCAGAGAGAGAGAGAGAGAGAGAG 1380
1322 CTGAGAAACTGGCTGAGCCCCAGAGAGTCCCCAGAGAGAGAGAGAGAGAGAGAG 1380
1381 TGAAGAGAGAGAGATGTGTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
1382 TGAAGAGAGAGAGATGTGTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
1441 CTGAAG 1500
1442 CTGAAG 1500
1501 CTCTCAGAGAAAGATCAAGATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
1502 CTCTCAGAGAAAGATCAAGATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
1561 GCTTAAAG 1620
1562 GCTTAAAG 1620
1621 AGCTGAG 1680
1622 AGCTGAG 1680
1681 GAGAGAGCTGTGCTGTCTCTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
1682 GAGAGAGCTGTGCTGTCTCTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
1741 GAG 1800
1742 GAG 1800
1801 AG 1860
1802 AG 1860
1861 TCCGAG 1920
1862 TCCGAG 1920
1921 TGTCTTCAGATGAG 1980
1922 TGTCTTCAGATGAG 1980
1981 AACAAG 2040
1982 AACAAG 2040
2041 TTTGTGTGATCATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2100
2042 TTTGTGTGATCATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2100
2101 GGGCAAG 2160
2102 GGGCAAG 2160
2161 CCGGAG 2220
2162 CCGGAG 2220
2221 CACCGAG 2280
2222 CACCGAG 2280
2281 AAAGATTAAGTCACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2340
2282 AAAGATTAAGTCACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2340
2341 CTAGGTGAG 2400
2342 CTAGGTGAG 2400

QY 2401 TTAAGAAATTCATCCCGGACGGCGGGAAGAAAGGGGACAGCGGGAAGCAAGAACGCCA 2460
DB 2401 TTAAGAAATTCATCCCGGACGGCGGGAAGAAAGGGGACAGCGGGAAGCAAGAACGCCA 2460
QY 2461 CTGTGGAAGACTCAGGCGCAGTGGAGATAAATGAGGACGACCCCTTAATGTCCCGCGGTG 2520
DB 2461 CTGTGGAAGACTCAGGCGCAGTGGAGATAAATGAGGACGACCCCTTAATGTCCCGCGGTG 2520
QY 2521 TGCTCTGTCTGAGTATAATGCAAGTGGAGAGGAGAAAGATGGAAGCCACGCGGGAATACGG 2580
DB 2521 TGCTCTGTCTGAGTATAATGCAAGTGGAGAGGAGAAAGATGGAAGCCACGCGGGAATACGG 2580
QY 2581 AGCTGCCCGACGCTCGGGGGCTGTGACGTGTCGAGGAGCTCAGTAAGACTCTGTGTC 2640
DB 2581 AGCTGCCCGACGCTCGGGGGCTGTGACGTGTCGAGGAGCTCAGTAAGACTCTGTGTC 2640
QY 2641 ACACCTGTGAGTGTGCGAGTCAATTTGATGGGACACGAGGCGAGTCACACAGTGTGCAAGAGCGGT 2700
DB 2641 ACACCTGTGAGTGTGCGAGTCAATTTGATGGGACACGAGGCGAGTCACACAGTGTGCAAGAGCGGT 2700
QY 2701 CTCCTTCGTGGATATCCGCTTCGTTAAACAGAACTCTTTGAACACACAGCGGGAGAACCCA 2760
DB 2701 CTCCTTCGTGGATATCCGCTTCGTTAAACAGAACTCTTTGAACACACAGCGGGAGAACCCA 2760
QY 2761 TGCCACCTGTGGAAGGTCACCTGAAAAGACATCATTTGCAAGAACTCTCTGTGCTCA 2820
DB 2761 TGCCACCTGTGGAAGGTCACCTGAAAAGACATCATTTGCAAGAACTCTCTGTGCTCA 2820
QY 2821 CCACAGCTTTACAGAGGTAAAGATGCCATGACGACATGCTCACCAGTGAAGTGGATT 2880
DB 2821 CCACAGCTTTACAGAGGTAAAGATGCCATGACGACATGCTCACCAGTGAAGTGGATT 2880
QY 2881 TCACCTCAGAAGCTGTGACAGCCACAGAGACCTCAGAGGCTCTCCGTACTGAAAGATT 2940
DB 2881 TCACCTCAGAAGCTGTGACAGCCACAGAGACCTCAGAGGCTCTCCGTACTGAAAGATT 2940
QY 2941 CCGAAGCATCGGGGGCGAAGAGACCAACAGACATGCTGTCGAGTTTCCAGCTGACTG 3000
DB 2941 CCGAAGCATCGGGGGCGAAGAGACCAACAGACATGCTGTCGAGTTTCCAGCTGACTG 3000
QY 3001 ACTCCCGACACACAGAGGAAGCCACCCAGTTCCAGGAGGTAGAGAGTGTGCTAG 3060
DB 3001 ACTCCCGACACACAGAGGAAGCCACCCAGTTCCAGGAGGTAGAGAGTGTGCTAG 3060
QY 3061 ATACAGAAGAGGAGCGCCAGACGAGGCCATCTCTCCAGCGCTTGACAGCAAGGTGA 3120
DB 3061 ATACAGAAGAGGAGCGCCAGACGAGGCCATCTCTCCAGCGCTTGACAGCAAGGTGA 3120
QY 3121 AAGAGAGTCCAGAGTCTGCAACCCAGACTGTGACAGAAACGGGGTCAAAAGCACTGG 3180
DB 3121 AAGAGAGTCCAGAGTCTGCAACCCAGACTGTGACAGAAACGGGGTCAAAAGCACTGG 3180
QY 3181 AGAAGGTTGAGGAGGTAGAGGAGCTCCGAGTGTGCTGAGCACTGTCAGAGGAGGAGACG 3240
DB 3181 AGAAGGTTGAGGAGGTAGAGGAGCTCCGAGTGTGCTGAGCACTGTCAGAGGAGGAGACG 3240
QY 3241 TTATGCCGAAGACCCGCTGCAGAGAGCTGGAGCTGAGCACTTTCGACAGGAGGCTCTGAGA 3300
DB 3241 TTATGCCGAAGACCCGCTGCAGAGAGCTGGAGCTGAGCACTTTCGACAGGAGGCTCTGAGA 3300
QY 3301 CTGGACAGGCTACTCCAGAGAGCTTTGAAGTTCCTGAAGTTCAGCGGAGATGTAGACCATG 3360
DB 3301 CTGGACAGGCTACTCCAGAGAGCTTTGAAGTTCCTGAAGTTCAGCGGAGATGTAGACCATG 3360
QY 3361 TCGCCACGCTCCAGGTTATCAAGCTCCAGAGCTGTATGGAACAGGCGGTGGCCCTGAGT 3420
DB 3361 TCGCCACGCTCCAGGTTATCAAGCTCCAGAGCTGTATGGAACAGGCGGTGGCCCTGAGT 3420
QY 3421 CATCCGAACCTTTGACAGACAGTGAACAAATGGAAGCACTCTCCCTAGCAGATTTCAGACA 3480
DB 3421 CATCCGAACCTTTGACAGACAGTGAACAAATGGAAGCACTCTCCCTAGCAGATTTCAGACA 3480

QY 3481 CTGCAGATGGGACACAGCAAGATGAAACCAATTCACAGCCAGGACAGTAAGCCACTGCGAG 3540
DB 3481 CTGCAGATGGGACACAGCAAGATGAAACCAATTCACAGCCAGGACAGTAAGCCACTGCGAG 3540
QY 3541 CTGTCAAGGAGTCAAGGCTCACAAGAAAGAGAGCGGCTACTGTCTCAGAAAGAGAGCCTT 3600
DB 3541 CTGTCAAGGAGTCAAGGCTCACAAGAAAGAGAGCGGCTACTGTCTCAGAAAGAGAGCCTT 3600
QY 3601 CGACACTACCTAATAATGTTTCCAGCCCAGGAAGAACATCGGGGAAGAGATG 3660
DB 3601 CGACACTACCTAATAATGTTTCCAGCCCAGGAAGAACATCGGGGAAGAGATG 3660
QY 3661 TTCTTTGAACCTACACAGCAAGAGCTTACTGCTCAGCGCTGCCCTTCTGCGAAAGACTG 3720
DB 3661 TTCTTTGAACCTACACAGCAAGAGCTTACTGCTCAGCGCTGCCCTTCTGCGAAAGACTG 3720
QY 3721 AGGTGGGTCAAGAGGTTGAGTTGACTGTTGGATGAGAGAAAGTCAAAAGAAAGACAGG 3780
DB 3721 AGGTGGGTCAAGAGGTTGAGTTGACTGTTGGATGAGAGAAAGTCAAAAGAAAGACAGG 3780
QY 3781 AGGTGTTTGTACACTCTGGAACCCACAGTCAAAAGGCTGCTGATGTGACATATGACAGTG 3840
DB 3781 AGGTGTTTGTACACTCTGGAACCCACAGTCAAAAGGCTGCTGATGTGACATATGACAGTG 3840
QY 3841 AAGTGTGGAGTGGCGGGTGTACAGAAAAGGAGTAGTCTGAAGTGCAGAGTCTTAGCC 3900
DB 3841 AAGTGTGGAGTGGCGGGTGTACAGAAAAGGAGTAGTCTGAAGTGCAGAGTCTTAGCC 3900
QY 3901 TGGAGAGGGAGAGATGGAACCTGACGTTGAAAGAGAAAGGGAGACAAAGCCAGAGC 3960
DB 3901 TGGAGAGGGAGAGATGGAACCTGACGTTGAAAGAGAAAGGGAGACAAAGCCAGAGC 3960
QY 3961 AAGTGTGAGTGAAGAGGTGAGCAGGAACACGCGCTCTCAGCATGAGGAACCTACGGGA 4020
DB 3961 AAGTGTGAGTGAAGAGGTGAGCAGGAACACGCGCTCTCAGCATGAGGAACCTACGGGA 4020
QY 4021 AGCCAGTCTGTGACACTGTGACATGCCAGCTCAGAGAGGGGAAAGGCACTGGGAAGCCTTG 4080
DB 4021 AGCCAGTCTGTGACACTGTGACATGCCAGCTCAGAGAGGGGAAAGGCACTGGGAAGCCTTG 4080
QY 4081 GAGGAAGCCCTTCTCTCCAGACCAAGACAAAGCAGGTTGCATAGAGGTTCAAGTTCAA 4140
DB 4081 GAGGAAGCCCTTCTCTCCAGACCAAGACAAAGCAGGTTGCATAGAGGTTCAAGTTCAA 4140
QY 4141 GCCTGGACACACAGTCACTCAAAACAGCAGAGAGCTGTGAAAAGGCTCATAGAAACGGTTG 4200
DB 4141 GCCTGGACACACAGTCACTCAAAACAGCAGAGAGCTGTGAAAAGGCTCATAGAAACGGTTG 4200
QY 4201 TGATTTTCAGAGACAGGTGAAAGTCCAGAGTGTGTAGGTGACACTTATTTACAGCTGAGA 4260
DB 4201 TGATTTTCAGAGACAGGTGAAAGTCCAGAGTGTGTAGGTGACACTTATTTACAGCTGAGA 4260
QY 4261 AGTCCTCTCAAGGGGTGGCCACTGACCTTTCAGCATGCGAGGAGACCGTTACCCCTGG 4320
DB 4261 AGTCCTCTCAAGGGGTGGCCACTGACCTTTCAGCATGCGAGGAGACCGTTACCCCTGG 4320
QY 4321 GGCTCAGTCTCAGGAGAGATCCATCCCAATCATAGTAACTCTCTGCTTCTGAAAGCACCC 4380
DB 4321 GGCTCAGTCTCAGGAGAGATCCATCCCAATCATAGTAACTCTCTGCTTCTGAAAGCACCC 4380
QY 4381 TACATCTGACCTTCAAGGAGAAATAAGCGCATCCAGAGAGAGCGATCAGAGGAAGAGG 4440
DB 4381 TACATCTGACCTTCAAGGAGAAATAAGCGCATCCAGAGAGAGCGATCAGAGGAAGAGG 4440
QY 4441 ACAGCCAGATGCTGCTGCTGATGCTGACGGCAGGAGAGATGACAGCAATGCAAAAAGTCC 4500
DB 4441 ACAGCCAGATGCTGCTGCTGATGCTGACGGCAGGAGAGATGACAGCAATGCAAAAAGTCC 4500
QY 4501 TCAAGGCTCAACCTGAGATCTCGGAACTTTGAGAGTGAAGCAACAGATTGCTCAAGC 4560
DB 4501 TCAAGGCTCAACCTGAGATCTCGGAACTTTGAGAGTGAAGCAACAGATTGCTCAAGC 4560
QY 4561 TCATTTCAGACAGCCGTTGACCAAGTTCGCACGTACAGAAAACAGCCCCGAAACTCATGCTT 4620


```

Db 4561 TCATTGAGACGCGGTTGACCACTTCGACGTACAGAAACAGCCCGAACTCATGCTT 4620
Qy 4621 ATGATTCACAGACCCAGGTTCTGTCATGCAAGCTTTGACAGAGGAGCCCAACAGATCT 4680
Db 4621 ATGATTCACAGACCCAGGTTCTGTCATGCAAGCTTTGACAGAGGAGCCCAACAGATCT 4680
Qy 4681 GGACAAAATGAAAGATGCAAGATGAAACACCCAGTCCCGCAGCCGAGAGAGATTCG 4740
Db 4681 GGACAAAATGAAAGATGCAAGATGAAACACCCAGTCCCGCAGAGGAGATTCG 4740
Qy 4741 AAGTCCTGACCGTTCTGGAGGATGGGCTCAGCTCGGAAATGCTTGCGCGCTTGCACT 4800
Db 4741 AAGTCCTGACCGTTCTGGAGGATGGGCTCAGCTCGGAAATGCTTGCGCGCTTGCACT 4800
Qy 4801 TGAAGCGCCGGTGTCAAAAGTAAAGCATTAAGAGAGTGCCTCTCAACCCAAAGATCCAA 4860
Db 4801 TGAAGCGCCGGTGTCAAAAGTAAAGCATTAAGAGAGTGCCTCTCAACCCAAAGATCCAA 4860
Qy 4861 AGAGCATGCTGCTGATGAGCTCCTCAGCTCAAGCTTGAGCCAGGACAGAGGAGCTGCT 4920
Db 4861 AGAGCATGCTGCTGATGAGCTCCTCAGCTCAAGCTTGAGCCAGGACAGAGGAGCTGCT 4920
Qy 4921 CTGGAACCTTAACCAAGATCCCGACACCAAGCTTGAGCCAGGACAGAGGAGCTGCT 4980
Db 4921 CTGGAACCTTAACCAAGATCCCGACACCAAGCTTGAGCCAGGACAGAGGAGCTGCT 4980
Qy 4981 ATCCCCCAAAAGTTGAGGTCCAGAGAAAGATGCTTACCAAGTCAAGTCAAAAGAAACA 5040
Db 4981 ATCCCCCAAAAGTTGAGGTCCAGAGAAAGATGCTTACCAAGTCAAGTCAAAAGAAACA 5040
Qy 5041 AGGCCAGGACAGAGAGAGAGCTGCAAGAGCCAAAGGAGAGCTGGCAAGATCTTAAGATG 5100
Db 5041 AGGCCAGGACAGAGAGAGAGCTGCAAGAGCCAAAGGAGAGCTGGCAAGATCTTAAGATG 5100
Qy 5101 TTAGTGTCTCATTTGATCAATCTGTAAGACCAAGATGTAAGAAACAGTCAAGAAACAAGATG 5160
Db 5101 TTAGTGTCTCATTTGATCAATCTGTAAGACCAAGATGTAAGAAACAGTCAAGAAACAAGATG 5160
Qy 5161 CTGCTGTTGGAGACTTGAAGACCAAGATTTCAAGGCCATGAGATCCAGAGAGAGGCGG 5220
Db 5161 CTGCTGTTGGAGACTTGAAGACCAAGATTTCAAGGCCATGAGATCCAGAGAGAGGCGG 5220
Qy 5221 TCCAAATGATTTCCACCAGTAGAGCACCAGCAATTCGAGGCTTCAATCGGAGCTTAA 5280
Db 5221 TCCAAATGATTTCCACCAGTAGAGCACCAGCAATTCGAGGCTTCAATCGGAGCTTAA 5280
Qy 5281 GCGAGCTAACATTTCTGCTTCAAGACGCTTGAATTTGCCCCCTTGAGTGCCTGCGTG 5340
Db 5281 GCGAGCTAACATTTCTGCTTCAAGACGCTTGAATTTGCCCCCTTGAGTGCCTGCGTG 5340
Qy 5341 TATTTCTAACATTTCTGCTTCAAGACGCTTGAATTTGCCCCCTTGAGTGCCTGCGTG 5400
Db 5341 TATTTCTAACATTTCTGCTTCAAGACGCTTGAATTTGCCCCCTTGAGTGCCTGCGTG 5400
Qy 5401 TATTTGGAATTAAGGTCTGCTTCAAGACGCTTGAATTTGCCCCCTTGAGTGCCTGCGTG 5460
Db 5401 TATTTGGAATTAAGGTCTGCTTCAAGACGCTTGAATTTGCCCCCTTGAGTGCCTGCGTG 5460
Qy 5461 ACTTTCGAAACGAGACATCTCTTATGATATTTATGATATTTATGATGCTGCTG 5520
Db 5461 ACTTTCGAAACGAGACATCTCTTATGATATTTATGATATTTATGATGCTGCTG 5520
Qy 5521 TCCGTGACCTATTTGATATTTTCTTAAGCTTTAAGCAATGCTTTTGTATTAAGCA 5580
Db 5521 TCCGTGACCTATTTGATATTTTCTTAAGCTTTAAGCAATGCTTTTGTATTAAGCA 5580
Qy 5581 TATATAACGGGTGTCAGCCATAGCAAGCTTTGAAAGCTTCAAGCTTCACTGTAAC 5640
Db 5581 TATATAACGGGTGTCAGCCATAGCAAGCTTTGAAAGCTTCAAGCTTCACTGTAAC 5640
Qy 5641 TGCAGCAACAGATTAATCTCTGTCGAGAGAGAGCAAGCTTTTAAAGTTAAGTAT 5700

```

```

Db 5641 TGCAGCAACAGATTAATCTCTGTCGAGAGAGAGCAAGTCTTTTAAAGTTACTGAT 5700
Qy 5701 GCTTAGATCTGTGGGCTTCTAGTCTCTGAAAGGTGTTTTCATAGCAGAGAGCT 5760
Db 5701 GCTTAGATCTGTGGGCTTCTAGTCTCTGAAAGGTGTTTTCATAGCAGAGAGCT 5760
Qy 5761 CAGAAATTAACCCCAATTTTGAACATCCAGATGTCCTCAATATTAACATGATTTTTC 5820
Db 5761 CAGAAATTAACCCCAATTTTGAACATCCAGATGTCCTCAATATTAACATGATTTTTC 5820
Qy 5821 CCCCCCTTTTCTAATCCAGTCCAGGTTGGAAAGATCTCTCTGTGTCAGATTAAGCC 5880
Db 5821 CCCCCCTTTTCTAATCCAGTCCAGGTTGGAAAGATCTCTCTGTGTCAGATTAAGCC 5880
Qy 5881 CTGTCTCTTAATGATGAGCAAAATGAGTGTGCTTAAGGCCATGAGATGTTTCTTAATG 5940
Db 5881 CTGTCTCTTAATGATGAGCAAAATGAGTGTGCTTAAGGCCATGAGATGTTTCTTAATG 5940
Qy 5941 AGAAGAAATCTGTTGATGATTTTGTGATTTGATGCTCTTAAGTCCAGACCAATTCATAT 6000
Db 5941 AGAAGAAATCTGTTGATGATTTTGTGATTTGATGCTCTTAAGTCCAGACCAATTCATAT 6000
Qy 6001 GCAGATCGAAGTGAATCTGTTCTTTTACAGATGATTTGATGATTAAGTGAATGCTG 6060
Db 6001 GCAGATCGAAGTGAATCTGTTCTTTTACAGATGATTTGATGATTAAGTGAATGCTG 6060
Qy 6061 TGTGTTATATCTGTGCCCCCTTTTAAAGCAATGTTGATTAATGCTTGTGATTAAT 6120
Db 6061 TGTGTTATATCTGTGCCCCCTTTTAAAGCAATGTTGATTAATGCTTGTGATTAAT 6120
Qy 6121 TGTGATTTGACACATGATTTTAAATTAACATATTTGACTAC 6160
Db 6121 TGTGATTTGACACATGATTTTAAATTAACATATTTGACTAC 6160

```

```

RESULT 2
US-09-902-432-1
; Sequence 1, Application US/09902432
; Patent No. US2002016002A1
; GENERAL INFORMATION:
; APPLICANT: Susan G. Gelman
; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
; FILE REFERENCE: A30558-A-FWC-A 070156,0597
; CURRENT APPLICATION NUMBER: US/09/902,432
; CURRENT FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 08/978,277
; PRIOR FILING DATE: 1997-11-25
; PRIOR APPLICATION NUMBER: 08/665,401
; PRIOR FILING DATE: 1996-06-18
; PRIOR APPLICATION NUMBER: 08/635,121
; PRIOR FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 1
; LENGTH: 5134
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-09-902-432-1

```

```

Query Match 78.4%; Score 4831.4; DB 9; Length 5134;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 5048; Conservative 0; Mismatches 76; Indels 31; Gaps 14;

Qy 1016 GGAAGAAAGACAGACCAAGCTTCGAGAGAGAGAGCCGCGAGAAACACAGACAGGCGG 1075
Db 1 GGAAGAAAGACAGACCAAGCTTCGAGAGAGAGAGCCGCGAGAAACACAGACAGGCGG 60

Qy 1076 GTTGTGACAGACTACGAGAAAGTGAAGTGCCTTTGGAAGACCGAGTGTGAGTCTGGA 1135
Db 61 GTTGTGACAGACTACGAGAAAGTGAAGTGCCTTTGGAAGACCGAGTGTGAGTCTGGA 120

Qy 1136 GGCATGTCAGAGAGAGAGTGTGCTCTTTGGCAACGGAAGTGTGATGAGAAATGGA 1195

```

Db	121	GGCATGCTCAGAGGAGAAGTGTCTCTCTTTGGCAACGGAAAGTGTGTGATGAGAAGATGGA	180
Qy	1196	AGCCCAACAAGAAAGTTGTTGCAGAGCTCCACGTGAGCAACCGTGGAGAAGACAGAGGAGGA	1255
Db	181	AGCCCAACAAGAAAGTTGTTGCAGAGCTCCACGTGAGCAACCGTGGAGAAGACAGAGGAGGA	240
Qy	1256	GCAGGAGGAGGAGAGAGGCTGAAGGGGGCGTGTGTGTAGAAAGAAACAGGAGAACTCTT	1315
Db	241	GCAGGAGGAGGAGAGAGGCTGAAGGGGGCGTGTGTGTAGAAAGAAACAGGAGAACTCTT	300
Qy	1316	GCCCCGTGAGAACTGCGCTGAGCCCCCAGAGAGTCCCCCAGGAAGCTGAGCCTGCTGAGGA	1375
Db	301	GCCCCCTTGAGAACTGCGTGAGCCCCCAGAGAGTCCCCCAGGAAGCTGAGCCTGCTGAGGA	360
Qy	1376	GCTGATCAAGAGCAGAGAGATGTGTCTCTCGAGGAGACACACTCAACTGACAGACCT	1435
Db	361	GCTGATCAAGAGCAGAGAGATGTGTCTCTCGAGGAGACACACTCAACTGACAGACCT	420
Qy	1436	AAGTCTTGAAGAGACGCTGCCAAAACCCAGAAAGCAATGTTCAGTGAGGTGGAGAT	1495
Db	421	AAGTCTTGAAGAGACGCTGCCAAAACCCAGAAAGCAATGTTCAGTGAGGTGGAGAT	480
Qy	1496	GCTGTCTCTCAGGAAGATCAAGGTAACAGGAAGTCCCTTGAGAAACTCTTCAGTAG	1555
Db	481	GCTGTCTCTCAGGAAGATCAAGGTAACAGGAAGTCCCTTGAGAAACTCTTCAGTAG	540
Qy	1556	CTCAGGCTTAAAGAAAGCTGTCTGGGAAGACAGAAAGGGGAAAACGAGAGGTGGGGAGA	1615
Db	541	CTCAGGCTTAAAGAAAGCTGTCTGGGAAGACAGAAAGGGGAAAACGAGAGGTGGGGAGA	600
Qy	1616	CGAAGCCTGGAGAATAACAACATTCACACCGAATCCCCAGAGAGTGTGTGATGAGCA	1675
Db	601	CGAAGCCTGGAGAATAACAACATTCACACCGAATCCCCAGAGAGTGTGTGATGAGCA	660
Qy	1676	GAAGGAGAGAGCTCTCGCTGTCCTCCCAAGAGAGCCTGAGAGACCAAGTGTCTGGAGAA	1735
Db	661	GAAGGAGAGAGCTCTCGCTGTCCTCCCAAGAGAGCCTGAGAGACCAAGTGTCTGGAGAA	720
Qy	1736	AGGCGCCTGGAAGCAACCCAGAGTGGGAAAGCTCAGGAAGCAACTCTTCCGATGGAGA	1795
Db	721	AGGCGCCTGGAAGCAACCCAGAGTGGGAAAGCTCAGGAAGCAACTCTTCCGATGGAGA	775
Qy	1796	GAGAAGAGAGAAGGATCACTCTCTGGGCATCTTTCAAAAAGATGTGACACCCAAAGAA	1855
Db	776	GAGAAGAGAGAAGGATCACTCTCTGGGCATCTTTCAAAAAGATGTGACACCCAAAGAA	835
Qy	1856	ACGGGTCGAAGACCTTCTGAGAGTGACAAGAGAGAAAGAGCTGGGAAGGTCGAAGAGCGC	1915
Db	836	AC--GGTCCGAAGACCTTCTGAGAGTGACAAGAGAGAAAGAGCTGGGAAGGTCGAAGAGCGC	894
Qy	1916	CACCTTGTCTCCACTGATAGACAGTGTCAAGAAATGCAAGATGAAGTCAAAACTGTGTGG	1975
Db	895	CACCTTGTCTCCACTGATAGACAGTGTCAAGAAATGCAAGATGAAGTCAAAACTGTGTGG	954
Qy	1976	TGAGGAACAAAAGCCAGAGAAACCAAGCGTAGGTGGATACTTCAGTGTCTTGGGAAGC	2035
Db	955	TGAGGAACAAAAGCCAGAGAAACCAAGCGTAGGTGGATACTTCAGTGTCTTGGGAAGC	1014
Qy	2036	ACTGATTTGTTCGGATCATCCAGAGAGACCAAGGAAGGATCTCTTCAGATGATGA	2095
Db	1015	ACTGATTTGTTCGGATCATCCAGAGAGACCAAGGAAGGATCTCTTCAGATGATGA	1071
Qy	2096	AGGAGGCCAAGGACACTCTGGGA--GGGGACAGTCACAGAGCAGAGGAGGCCAGCAAGACA	2154
Db	1072	AGGAGGCCAAGGACACTCTGGGAAGGGGACAGTCACAGAGCAGAGGAGGCCAGCAAGACA	1131
Qy	2155	AAGAAGCCGAAACAGACGCTGTTCTCTGCCACACCCAGAGCAGGACCAAGCGCAAGGAA	2214
Db	1132	AAGAAGCC--GAAACAGACGCTGTTCTCTGCCACACCCAGAGCAGGACCAAGCGCAAGGAA	1190
Qy	2215	GTTCTCACCCGAGCAGCGGGAAGCCCTTCGAAAGGGGAAGGTTCTTCACTTGGGAGT	2274

QY 3346 CAGATGTAGACCATGTGTCCACGTCGAGTTATTCAACTCCAGAGCTGATGTGAACAG 3405
Db 2328 CAGATGTAGACCATGTGTCCACGTCGAGTTATTCAACTCCAGAGCTGATGTGAACAG 2387
QY 3406 CCGTGGCCCCGTGAGTCATCCGAAACCTTGACAGACAGTGAACAAATGGAAGCACTCCCT 3465
Db 2388 CCGTGGCCCCGTGAGTCATCCGAAACCTTGACAGACAGTGAACAAATGGAAGCACTCCCT 2447
QY 3466 TAGCAGATTGACAGACTGTCAGATGGGACACAGCAAGATGAACCAATTGACAGCCAGACA 3525
Db 2448 TAGCAGATTGACAGACTGTCAGATGGGACACAGCAAGATGAACCAATTGACAGCCAGACA 2507
QY 3526 GTAAAGCCACTGCACTGTCAAGGAGTCAAGGTCAAGAAAGAAAGCGGCTACTGCTC 3585
Db 2508 GTAAAGCCACTGCACTGTCAAGGAGTCAAGGTCAAGAAAGAAAGCGGCTACTGCTC 2567
QY 3566 AGAAAGAGAGCCTTCGACACTAATTAATGTTCCAGCCAGGAAGAACATGGGGAG 3645
Db 2568 AGAAAGAGAGCCTTCGACACTAATTAATGTTCCAGCCAGGAAGAACATGGGGAG 2627
QY 3646 AACCGAAGAGATGTTCTTGAACCTTACACAGCAAGCTTACTGCTCAGCCGTCGCCG 3705
Db 2628 AACCGAAGAGATGTTCTTGAACCTTACACAGCAAGCTTACTGCTCAGCCGTCGCCG 2687
QY 3706 TTTCTGGCAAAAGACTGAGGTGGGTCAAGAGGGTGAAGGTTGACTGTTGGATGGAGAAAAG 3765
Db 2688 TCTGGCAAAAGACTGAGGTGGGTCAAGAGGGTGAAGGTTGACTGTTGGATGGAGAAAAG 2747
QY 3766 TCAAGAAAGAACAGAGAGTGTGTTGTCACCTCGGACCCAAAGCTCAAAAGGCTGCTGATG 3825
Db 2748 TCAAGAAAGAACAGAGAGTGTGTTGTCACCTCGGACCCAAAGCTCAAAAGGCTGCTGATG 2807
QY 3826 TGACATATGACAGTGAAGTATGAGAGTGGCCGGGTGTGTCAGAGAAAAGAGACTACTGAAG 3885
Db 2808 TGACATATGACAGTGAAGTATGAGAGTGGCCGGGTGTGTCAGAGAAAAGAGACTACTGAAG 2867
QY 3886 TGCAGAGTCTTATGCTTGGAGAGGAGAGATGGAACCTGACCTTGAAGAGAGAGAGAGAG 3945
Db 2868 TGCAGAGTCTTATGCTTGGAGAGGAGAGATGGAACCTGACCTTGAAGAGAGAGAGAGAG 2927
QY 3946 AGACAAAGCCAGAGAGAGTGAAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4005
Db 2928 AGACAAAGCCAGAGAGAGTGAAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2987
QY 4006 AAGGAACCTACGGGAGAGCCAGTCTGACACTTGACACTGAGCTCAGAGAGGGGAGAG 4065
Db 2988 AAGGAACCTACGGGAGAGCCAGTCTGACACTTGACACTGAGCTCAGAGAGGGGAGAG 3047
QY 4066 CACTGGGAGAGCCTTGGAGAGAGCCCTTCTCTCCAGAGCCAAAGCAAGAGAGAGAGAGAG 4125
Db 3048 CACTGGGAGAGCCTTGGAGAGAGCCCTTCTCTCCAGAGCCAAAGCAAGAGAGAGAGAGAG 3107
QY 4126 AGGTTCAAGTTTAAAGAGCTGAGACAAACAGTCACTCAACACAGAGAGAGTGGGAGAGAG 4185
Db 3108 AGGTTCAAGTTTAAAGAGCTGAGACAAACAGTCACTCAACACAGAGAGAGTGGGAGAGAG 3167
QY 4186 TCATAGAAACGGTGTGATTTTCAAGAGACAGTGAAGAGTCCAGAGTGTAGAGTGCACACT 4245
Db 3168 TCATAGAAACGGTGTGATTTTCAAGAGACAGTGAAGAGTCCAGAGTGTGTAGAGTGCACACT 3227
QY 4246 TATTTACAGCTGAGAGAGTCTCTGCAACGGGTGGCCACTGCACTCTTCAGCATGCAAGAG 4305
Db 3228 TATTTACAGCTGAGAGAGTCTCTCTGCAACGGGTGGCCACTGCACTCTTCAGCATGCAAGAG 3287
QY 4306 ACAGGGTACCCCTGGGGCTGAGTCTCAGGCAAGATCCATCCCAATCAATAGTAATCTCTG 4365
Db 3288 ACAGGGTACCCCTGGGGCTGAGTCTCAGGCAAGATCCATCCCAATCAATAGTAATCTCTG 3347
QY 4366 CTCCTGAAAGGACCCCTACATCCCTGACCTTCAAGAGAGAAATTAAGCGCATCCAGAGAGAG 4425
Db 3348 CTCCTGAAAGGACCCCTACATCCCTGACCTTCAAGAGAGAAATTAAGCGCATCCAGAGAGAG 3407

QY 4426 GATCAGAGAAAGAGACAAAGCCAGATGCTGTGCTCTGATGTCAGACGGCAAGAGATACAG 4485
Db 3408 GATCAGAGAAAGAGACAAAGCCAGATGCTGTGCTCTGATGTCAGACGGCAAGAGATACAG 3467
QY 4486 CAATCGAAAAAGTCTCAAGGCTGAAACCTGAGATCTTGAACCTTTAGAGATTAAGACAACA 4545
Db 3468 CAATCGAAAAAGTCTCAAGGCTGAAACCTGAGATCTTGAACCTTTAGAGATTAAGACAACA 3527
QY 4546 AGATTGTCTGAACAGTCAATTCAGACAGCCGTTGACAGTTTGCACCTGTAACAGAAACACCC 4605
Db 3528 AGATTGTCTGAACAGTCAATTCAGACAGCCGTTGACAGTTTGCACCTGTAACAGAAACACCC 3587
QY 4606 CCGAACTCATGCTTATGATTCACAGACCAAGTTCCTGCAATGCAAGGCTTGAACAGCAGAG 4665
Db 3588 CCGAACTCATGCTTATGATTCACAGACCAAGTTCCTGCAATGCGCTTGAACAGCAGAG 3647
QY 4666 AGCCCAACAGTGTGACAAAAATGAAAGATGCAAGATGAACACCCAGTGTCCGAGC 4725
Db 3648 AGCCCAACAGTGTGACAAAAATGAAAGTTCAGAGATGAACACCCAGTGTCCGAGC 3707
QY 4726 CCAAGAGAGATTGCAAGTCTGACCGTTCGAGAGCATGGGCTCAGCTTGGAAATGCT 4785
Db 3708 CCAAGAGAGATTGCAAGTCTGACCGTTCGAGAGCATGGGCTCAG--CTCGAAATGCT 3765
QY 4786 TGCAGGCTGTCAGTTGTAAGAGCGCGGCTGCAAGTGAAGCATGGAAGAGTGCCTCTCA 4845
Db 3766 TGCAGGCTGTCAGTTGTAAGAGCGCGGCTGCAAGTGAAGCATGGAAGAGTGCCTCTCA 3825
QY 4846 ACCCAAGATCCAAAGAGAGCATGCTGTGATGACCTCAGCTTCAAGCTTAAAGCCAGG 4905
Db 3826 ACCCAAGATCCAAAGAGAGCATGCTGTGATGACCTCAGCTTCAAGCTTAAAGCCAGG 3884
QY 4906 CAGAGGCAATGCTCTGAGAAACCTTAACCAAGAAATCCCAAGACCAACAGGACCAAGC 4965
Db 3885 CAGAGGCA--GTGTGAGAAACCTTAACCAAGAAATCCCAAGACCAACAGGACCAAGC 3941
QY 4966 TAAACAGAGAGAGGCGATCCCAAAAGTTGAGGTCCAGAGAGAGAAATGCTCAACAGT 5025
Db 3942 TAAACAGAGAGGCGATGCTGAGAGTTGAGGTCCAGAGAGAGAAATGCTCAACAGT 4000
QY 5026 CAGTCAAAAGAGACAAAGGCCAGGACAGAGAGACCTGACAGAGCCAAAGGAGAGCCTGG 5085
Db 4001 CAGTCAAAAGAGACAAAGGCCAGGACAGAGAGACCTGACAGAGCCAAAGGAGAGCCTGG 4060
QY 5086 CAGAACTCTAAGATGTTAGTGTCTGCTCATCTGTGTAACCAAGATGTGAAGAAACAG 5145
Db 4061 CAGAACTCTAAGATGTTAGTGTCTGCTCATCTGTGTAACCAAGATGTGAAGAAACAG 4120
QY 5146 TCACAGAACAAAGATGCTGTGAGAGCCTTGAAGCCAAAGATTCAGAGCCCATGAGATC 5205
Db 4121 TCACAGAACAAAGATGCTGTGAGAGCCTTGTGGAGCTTG-CAACAAAGATTCAGAGCCCATGAGATC 4179
QY 5206 CAGAGAGAGGAGCCGTCATATGATTTTCCACCAGTAGAGCAACCCGACAAATTTGAGGCT 5265
Db 4180 CAGAGAGAGGAGCCGTCATATGATTTTCCAAATTTCTGTTTAAAGCTGCTTTGATTTGCCCC 4239
QY 5266 TCATGGSAGCTAGAGCCAGCTTAACATTTCTCGTTTCAAGCTGCTTTGATTTGCCCC 5325
Db 4240 TCATGGSAGCTAGAGCCAGCTTAACATTTCTCGTTTCAAGCTGCTTTGATTTGCCCC 4299
QY 5326 TTGATGCGGTCCGTGATTTTCAACATTTCTCGTTTCAAGCTGCTTTGATTTGCCCC 5385
Db 4300 TTGATGCGGTCCGTGATTTTCAACATTTCTCGTTTCAAGCTGCTTTGATTTGCCCC 4359
QY 5386 TTGATGCGGTCCGTGATTTTCAAGCTGCTGCTTCAACCTGGAACCAATTTCT 5445
Db 4360 TTGATGCGGTCCGTGATTTTCAAGCTGCTGCTTCAACCTGGAACCAATTTCT 4419
QY 5446 GGCATACCTAAGTTCACATTTCTCAACCTGAGACATCTCTTTATGATTTATGATGT 5505
Db 4420 GGCATACCTAAGTTCACATTTCTCAACCTGAGACATCTCTTTATGATTTATGATGT 4479
QY 5506 TTTATGTAGTCTCTCTCTGTAACCTATTTGATTTTCTTCAAGTTTAAGCAATGCT 5565

QY 849 AGCAGCCAGTCAACGATGAGCAACATCTTCTTCAAGAACTTCTCACTCAGCGTTGG 908
DB 1041 ACTAGTCCCGTGACGACGATGAACAGATCAACCTTCAAAAAATTTCTCACTCAAGGTTGG 1100
QY 909 GCCGCTGCGCAAGAAACCAAGCTTCAAGAAATCAAAAGAGATGATCTGGAACTGCC 968
DB 1101 GCCGCTGCGCAAGAAACCAAGCTTCAAGAAACCGAAAGAGATGAAAGTGAAGCTTCA 1160
QY 969 GAGAAAGAAAGAGACAGAGGACAGAAAAAGTGAACGAGAAAGAAAGAAAAAGACAGAG 1028
DB 1161 GAGAAAGAAAGAGACAGAGGACAGAAAAAGTGAACAGAAAGAAAGACGAAAGGACAGAG 1220
QY 1029 CCAGCTTCGAGGA-----CGAGAGCCCGCCAGAAAC 1061
DB 1221 GTTGCTCCGAGAAACAGACCCCTCCGAGCAAGCCCAAGAGACCGGCAAGAAAGT 1280
QY 1062 ACAGACCAAGGCGAGGTTCAGACAGACTACGAGAGGTGAGAGCTCTTTGGAAAGCCAG 1121
DB 1281 GCCACGAGCCCGGTTATCACTGAAATGAGAAAGTTGAGCTGCCCTCAGAGAGACAA 1340
QY 1122 GTTGCTGACCTGAGAGCATCGTCAAGAGAGAGTGTCTCTTTGGCAAACGAGAGTGT 1181
DB 1341 GTCAATGGCTGACAGGACCTTCTGAAAGAAACCTGCTCCGTTGGCGACAGAAAGTGT 1400
QY 1182 GATGAAAGATGAAGCCCAACAGAA---GTTGTTGACAGAGTCCAGTGAACCCGTG 1238
DB 1401 GATGAAAGATGAAGTCCCAACAGAAAGAGTGTGTGCGGAGAGTCCACCTGACACCGTG 1460
QY 1239 GAGAAAGACAGAGAGAGAGGAGAGAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1298
DB 1461 GAGGAGAGAACCGAGAGAGAGAA-----ACGAGAGTGA 1496
QY 1299 GGAACAGAGAAATCCTTGGCCCTGAGAAATCGCTGAGCCCGAGAGAGTCCCGAGAA 1358
DB 1497 GAAACAGAGAGAGTGTGTGCACTGAAGATTTGGTTGAATGATGACAGAACTCAGAA 1556
QY 1359 GCTGAGCTCTGAGAGAGTGAATGAAGACAGAGATGTGTCTCTGAGAGAGACCCAC 1418
DB 1557 GCGGAACCTGACAGAGAGCTGTGAAAGCTCAAGAAACCTGTGTTCGAGAGAGACCT 1616
QY 1419 ACTCAACTGACAGACCTAAGTCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1478
DB 1617 AACACAGGAGCTGACCTCAGTCTGATGAGAGAGTGTCTCAAAACCCCGAGAGCGTT 1676
QY 1479 GTCAAGTGAAGTGAAGTGTCTCTCAGAGAAAGATCAAGGTACAGGAGAGTCCCTTG 1538
DB 1677 GTGAGTGAAGTGAAGTGTCTCTCAGAGAGAGATCAAGGTGAAGTGAAGTGAAGTGA 1736
QY 1539 AAGAAATCTTCAAGTGAAGTCAAGGCTTTAAAGAGTGTCTGAGAGAGAGAGAGAGAG 1598
DB 1737 AAGAAAGCTTTTACAGCACTGCTTAAAAAGCTTTCTGAGAAAGAAACAGAAAGGAAA 1796
QY 1599 CGAGAGAGTGGGAG 1658
DB 1797 AAGAGG---GAGAGAGACAGAGAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1853
QY 1659 GAGAGTGTGATGAG 1718
DB 1854 GACAG 1913
QY 1719 ACACAGTGTCTGAG 1778
DB 1914 ATCAAGTGTCTGAG 1973
QY 1779 ACTACTTCGATGAG 1838
DB 1974 GCTACTTCGATGAG 2033
QY 1839 ATGAGTGAACCCCAAGAAAG 1898
DB 2034 ATGAGTGAACCCCAAGAAAG 2093

QY 1899 GAGAAAGTCAAG 1958
DB 2094 GACAAAGTCAAG 2153
QY 1959 GAAGTCAAAACCTGTTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2018
DB 2154 GAAATGAAG 2213
QY 2019 TCAAGTGTCTGAG 2078
DB 2214 TCAGTATCTTGGAG 2273
QY 2079 TCTCTTCAAGTATGAG 2138
DB 2274 TCTCTTCTGATGAG 2333
QY 2139 GAGGCGACAAAGACAAAG 2198
DB 2334 GAGGCGGAG 2393
QY 2199 GACCAAGGCGAG 2258
DB 2394 GATCCAGGCGAG 2453
QY 2259 GTCTCACTTGGAG 2318
DB 2454 GTTTCACCTTGGAG 2513
QY 2319 CTGGAAG 2366
DB 2514 CTGGAAG 2573
QY 2367 ATCGAACCGAGATGAG 2426
DB 2574 ACTGAACCGGTTAAAG 2633
QY 2427 AAGAAAG 2486
DB 2634 AAGAAAG 2693
QY 2487 ATAAATGAG 2546
DB 2694 GCGAAG 2753
QY 2547 GAGAGGAG 2597
DB 2754 GAAAGGAG 2813
QY 2598 GGGGCTGTGTGAG 2657
DB 2814 GCGACACAG 2873
QY 2658 GTCAATGATGAG 2717
DB 2874 GTCGCTGACGGGAG 2933
QY 2718 GCTTCCGTAACAG 2777
DB 2934 GCTTATGATGAG 2993
QY 2778 GTCACTGAAAAAGACATGATGAG 2834
DB 2994 GTATTGAG 3053
QY 2835 GAGGTAAGATGAG 2894
DB 3054 GAGAAAG 3113
QY 2895 GTGACAG 2954
DB 3114 GTGACAG 3173
QY 2955 GCGGAAG 3014

[illegible]

Db	4254	GAGAAGCTTGAGCAGCAAGACAGCTGTTACCGTATCTGAAGAGGTCGTAAGCAGCGCTCCTC	4313
Qy	4029	CTGACACTTGCATGCCAGCTCAGAGAGGGGGAAGGCACCTGGGAAGCCCTTTGGAGGAAGC	4088
Db	4314	CAGACAGTGAATGTGCCATCATAGATGGGCNAAAGGAAGTCAGCAGTTTTCGAAGGAAGC	4373
Qy	4089	CTCTTCTC---TCCCAGACCAAGCAAGCAGAGTTGCGATAGAGGTTCAAGTTCAAAGCCTG	4145
Db	4374	CCTCCTCCCTGCCTTAGTCAAGAGGAGGAGTAGTGCACCAAAATTCAAAGTTTCAGAGTCT	4433
Qy	4146	GACACACAGTCACTCAAAACAGCAGAAAGCTGTGGAAAA-----GGTCATAGAACGGTT	4199
Db	4434	GAGGATCATTTCACTTAAACAGCGGCTGCAGAGGAGAAAAAGGTCTTAGGAAGAACTGCC	4493
Qy	4200	GTGATTTTCAGAGACAGGTGAAAGTCACAGAGTGTGTAGGTGCACACCTTATTACCAAGCTGAG	4259
Db	4494	AACATTTTAGAAACAGGTGAAACGTTTGGAGCTGCAGGTGCACATTTAGTTTCTGGAAGAG	4553
Qy	4260	AAGTCTCTGAAACGGGTGGCACA CTGGACTCTTTCAGCATGCAGAGGACACGGTACCCCTG	4319
Db	4554	AAATCTCTGAAAAAATGAAGACTTTGCGCTCATCCAGGGGAAGATGCTGTGCCACACA	4613
Qy	4320	GGGCTCAGTCTCAGGCAGAACTCCATCCCAATCAGTAGTAACTTCCTGCTCCTGAAAGCACC	4379
Db	4614	GGGCGGACTGTGAGGCAAAATCGACACAGTGTAGTAGTATCTGCTACTACCAAGAAAGGC	4673
Qy	4380	CTACATCTCGACCTACAAGGAGAAAAATAAGCGCATCCACAGAGAGCGATCAGAGGAAGAG	4439
Db	4674	TTAAGTTCGACCTGGAAGGAGGAAAAACCATCACTGAAGTGAAGTCAGATGAAAGTC	4733
Qy	4440	GACAAGCCAGATGCTGGTCTGTGATGCTGACGCAAGGAGAGTACAGCAATCGAAAAAGTC	4499
Db	4734	GATGAGCAGGTTGCTTGC-----AGGAGGTCAAAGTGAGTGTAGCAATTG---AGGAT	4784
Qy	4500	CTCAAGGCTGAACCTGAGATCTCTGGAACCTTGAGAGTAAAGCAACAGATTTGCTGTAAC	4559
Db	4785	TTAGAGCTTGAAAAATGGGATTTTGGAACTTTGAGACCAAAAGCAGTAAACTTTGTCCAAAAC	4844
Qy	4560	GTCAATTCAGACAGCCGTTGACCAAGTTGCGCACGCTAC---AGAAACAGCCCGGAAACTCAT	4616
Db	4845	ATCATCAGACAGCCGTTGACCAAGTTTGTACGTACAGAAAGAAACAGCCACCGAAATGTTG	4904
Qy	4617	GCTTATGATTACAGACCCAGGTTCTC-----CTGCATGCAGGCTTGACAG	4660
Db	4905	ACGTCTGAGTTACAGACACAAGCTCAGCTGATAAAAGCTGACAGCCAGACGCTGGACAG	4964
Qy	4661	CAGGAGCCCAACAGATGCTGGACAAAAATGAAGATGCCAAGATGAAGCAACCCAGTGCC	4720
Db	4965	GAACCGAGAAAGAGGAGGAAACCTCAGGCTCTGCACAGGATGAACACCAATTAAT	5024
Qy	4721	GCAGCCCA-----GAGAGACTTGCAAGTTCCTGACCGTTCTGGA	4759
Db	5025	TCAGCCAAAGAGAGTCAGAGTCAACCGCAGTGGGCAAGCACATTTCTGATTTTCCAAA	5084
Qy	4760	GGCATGGGCTCAGCCCTCGGAAATGCTTGC CGCGCTTGCGATTTGAAAGCGCGGTGTCAA-	4818
Db	5085	GACATGAGTGAACCTCAGAAAGACCATGACTGTTGAGGTAGAAAGTTTCCACTGTAAAT	5144
Qy	4819	-----	4818
Db	5145	GATCAGCAGCTGGAAGAGGTCTGCTCCCATCTGAGGAAGAGGAGGTGGAGCTGGAACA	5204
Qy	4819	-----AGTAAGCATTTGAGAGCTGCCT	4840
Db	5205	AAGTCTGTGCCAAGAGATGATGGTATGCTTTGTAGCAGAAAGAAATAGAGAAAGTCACTA	5264
Qy	4841	CCTCAACCCAAAGA--TCCAAAAGGAGCATGCTGCTGATGCCCTCAGCTTCCAAAGCTTA	4898
Db	5265	GTTGACCGAAGAGATGAAAAAGGTGATGATGTTGATGACCTGTAAACCCAGAACTCA	5324
Qy	4899	GCCCAGCAGAGGCCAGTGCCTCTTGAAACCTTAACCAAGAAATCCCAGACACCAACGGA	4958
Db	5325	GCCTGCTGATACTGATGCTCCTCAGAGGCTTAACCAAGAGTCCCAGATACAATATGGA	5384

Db 522 GACTCTGAAGATGTAGCGAAAGAGACTCCGATTAAGAGATGGCTACTTAAGTCAAGCGTT 581
Qy 369 GTTGAAGATATCAAAAGGATGGCAGAGGAGACATCAGAAATAATTTGAACAGATCCCT 428
Db 582 GTTCACGACATCACAGATGATGGCAGGAGGAGA---ACCGAAATATCGAACAGATTCCT 638
Qy 429 GCTTCAGAAACAATGTGAAGAAATGGTACAGCGCTGTGATGCCAGGCTTAATGATGTT 488
Db 639 TCTTCAGAAAGCAATTTAGAAAGAGCTAACCAACCCCACTGATGCCAGGCTTAATGATATT 698
Qy 489 GGCTTCAAGAAAGTATTTAAATTTGTTGTTTAAATTTCAAGGTGAAGAGATAAATAAT 548
Db 699 GGATTTAAGAAAGTGTTTAAGTTTGTGGCTTTAAATTTCACTGTGAAGAAAGGATAAGACA 758
Qy 549 GAAAGTCAGATCTCTCAACTACTCACTGTCAAGAGGATGAAGGCGAAGGGGCGAGAA 608
Db 759 GAGAAGCCTGACACTGTCCAGCTACTCACTGTGAAGAAAGATGAAGGGGAGGAGCACA 818
Qy 609 GCCTCTGCGAGCTGAGACACACAGAGCCCAAGTGTGAGACTGCCGTGCGAGAGTCA 668
Db 819 -----GGGGCTGGCGACACAGAGACCCCAAGCCTTGGGGCT-----GGAGAAAGCA 863
Qy 669 GCATCCAAAGAAAGTGAAGTCAAGCAATCCACAGAGAGCAAGAGGACCCCTGAAGCAA 728
Db 864 GCATCCAAAGAAAGCGAAACCCCAACAATCTACAGAGAAACCCGAAGAGACCCCTGAAGCGT 923
Qy 729 GAACAGAGCAGCAGAGAAATCCCTCTCAAGCCGAATCTGATCAAGCGGCTGAGAGAA 788
Db 924 GAGCAAGCCACGAGAAATTTCTCCCGCCCGAATCTGGCCAAAGCAGT---GGAGGAA 980
Qy 789 GCCAAAGATGAAGAGAGAAAGAAACAAGAGAAAGAGCCCAAGTCCCCAGATCCCCG 848
Db 981 TGCAAGAGGAAGAGAGAGAAACAAGAAAGAAACCTAGCAAGTCTGCAGAAATCTCCG 1040
Qy 849 AGCAGCCAGTCAACAGTGAAGCAACATCTTCTTCAAGAGTCTTCACTCAAGGTTGG 908
Db 1041 ACTAGTCCCGTGAACAGTGAACAGGATCAACCTTCAAAAAATTTCTCACTCAAGGTTGG 1100
Qy 909 GCCCGCTGGCCAAAGAGACAGCTTCAAGAAATCAAAAGAGGATGATCTGGAATCTGCC 968
Db 1101 GCCCGCTGGCCAAAGAGACCAAGTTCAGGAAGCGGAAGAGATGAAGTGAAGCTTCA 1160
Qy 969 GAGAAGAAAGAGAGCAGAGCGCAAGAAAGTGAAGAGAGAAAGAAAGAAAGAGAGAG 1028
Db 1161 GAGAAGAAAG 1220
Qy 1029 CCAGCCTCGAGGA-----GGAGGAGCCGCGAGAGAC 1061
Db 1221 GTTCCTCCGAGAAATGACCGCTCCGAGCAAGCCCAACCCACAGAGGCGCGCAGAAAGT 1280
Qy 1062 ACAGACCAAGGCGAGGTTGTCAAGCAGACTACGAGAAAGGTGAGCTGCCCTTTGGAAGACAG 1121
Db 1281 GCCACAGGCCCGGTTATCAGCTGAATATGAGAAAGTGTGAGCTGCCCTCAGAGAGCAA 1340
Qy 1122 GTTGGTACCTGGAGGCAATCTGAGAGAGAGAGTGTCTCTTTGGCAACGGAAAGTGTTT 1181
Db 1341 GTCAAGTGGCTCGCAGGAGACCTTCTGAAGAGAAACCTGTCTCGTTGGCGACAGAAGTGT 1400
Qy 1182 GATCAGAGATGGAAGCCCAACAGAA---GTTGTTGAGAGGTTCCAGTGAAGCAGCTG 1238
Db 1401 GATCAGAAATAGAAAGTCCACCAAGAGAGGTTGTGGCCGAAGTCCAGCTCAGACCCGTG 1460
Qy 1239 GAGAAGACAGAGGAGGAGGAGGAGGAGGAGGAGGCTGAAGGGGCGTGGTGTAGAA 1298
Db 1461 GAGGAGAGACCGAGAGCAGAAA-----ACGAGGTGGAA 1496
Qy 1299 GGAACAGAGAAATCTTGGCCCTCTGAGAACTGGCTGAGCCCGCAGGAGGTCCCCAGAGAA 1358
Db 1497 GAAACAGCAGGCTGTGTGCCAGCTGAAGAAATTTGTTGGAATGGATGCAGAACCTCAGGAA 1556
Qy 1359 GCTGAGCCTGCTGAGGAGCTGATGAAGAGCAGAGAGATGTGTCTCTGAGGAGAGACCAC 1418

Db 1557 GCCGAACCTGCCAAGGAGCTGGTGAAGCTCAAAGAAACGTTGTGTTCCGGAGAGGACCCT 1616
Qy 1419 ACTCAACTGACAGACCTTAAGTCTCTGAAGAGAGACGCTGCCCAAAACCCAGAGAGGCATT 1478
Db 1617 ACAGAGGAGCTGACCTCAGTCTGATGAGAGGCTGCTGCTCCAAACCCCGAAGGCGTT 1676
Qy 1479 GTCAAGTGAAGTGGAGATGCTCTCTCAGAAAGAAATCAAGGTACAGGGAAGTCCCTTG 1538
Db 1677 GTGAGTGAAGTGGAAATGCTGTCTCATCAAGGAGAGAAATGAAGGTGCAGGGAAGTCCACTA 1736
Qy 1539 AAGAAACTCTTCTAGTACTCAGGCTTAAAGAGCTGTCTGGAAGAAAGACAGAGGGGAAA 1598
Db 1737 AGAAGCTTTTACCAGACTGGCTTAAAGAGCTTCTGAAAGAAACAGAAAGGAAA 1796
Qy 1599 CGAGGAGTGGGGAGACGAGAGCCTTGGAGATACCAACATTCACCCGAATCCCCA 1658
Db 1797 AGAGGA---GGAGGAGACGAGGAATCAGGGGAGCACACTCAGGTTCCAGCCGATTCTCCG 1853
Qy 1659 GAGAGTCTGATGAGCAGAAAGGAGAGAGCTCTGCGTCTGCCCGAGGAGGACCTGAGGAG 1718
Db 1854 GACAGCCAGGAGGAGCAAAAGGGCGAGAGCTCTGCTTCATCTCCCTGAGGAGCCCGAGGAG 1913
Qy 1719 ACCACGTGTCTGGAGAAAGGCGCGCTGGAAGCACCCAGATGGGGAAGCTTGAGGAAGGA 1778
Db 1914 ATCAGTGTCTGAAAGAGGCTTAGCCAGGTGCAGCAGATGGGGAAGCTGAAAGAAAGGA 1973
Qy 1779 ACTACTTCCGATGGAGAGAAAGAGAGAGGATCATCTCCCTGGGCATCTTCAAAAAG 1838
Db 1974 GCTACTTCCGATGGAGAGAAAGAGAGAGGTTGCTCACTCCCTGGGCATCATTCAAAAG 2033
Qy 1839 ATGTTGACACCCAAAGAAACGGGTCCGAAGACTTCTGAGAGTGAAGAGGAGAGAGAGCTG 1898
Db 2034 ATGTTGACGCGCCAAAGAGCGTTTAGACGCGCTTCGGAAGAGTGAATAAGAAAGATGAGCTG 2093
Qy 1899 GAGAGGTCAAGAGCGCCACTTGTCTCTCACTGATAGCAGCTGTGAGAAATCCAGAT 1958
Db 2094 GACAAGTCAAGAGCGCTACTCTTCTCCAGCAGAGCAGCGCTCTGAAATCAAGAA 2153
Qy 1959 GAAGTCAAAACTGTGTTGGTGAAGAACAAAGCCAGAGAGAACAAAGCGTAGGGTGGATACT 2018
Db 2154 GAAATGAAGGGAGCGTTGGAAGAGCCAAAGCCGGAAGAACCAAGCGCAGGTGATACC 2213
Qy 2019 TCAGTGTCTTTGGAGACACTGATTTGTGTCGATCATCTCAAGAAAGAGAGAGAGAGAGCA 2078
Db 2214 TCAGTATCTTTGGAGAGCTTTAAATTTGTGGGATCATCTCAAGAAAGAGAGAGAGAGG 2273
Qy 2079 TCCTCTTCAGATGATGAAGGAGGCGCAAGGACACTGGAGGAGGACAGTCAAGAGCAGAG 2138
Db 2274 TCCTCTTCTGATGAGGAAGGGGACCAAAAGCAATGGAGAGAGACCACCAAGAAAGCTGAT 2333
Qy 2139 GAGGCCAGCAAGAACAAAGAGCGGAACAGACGCTCTTCTGCCAGACCCAGGAGCAG 2198
Db 2334 GAGGCCGGAAGAACAAAGAGACGGGACAGACGGGATCTTGTGTTTCCCAAGAACAT 2393
Qy 2199 GACCAAGCGCAAGAAAGTTCTCAACCCAGCAGCGGGAAGCCCTTCCGAAGGGAAGGT 2258
Db 2394 GATCCAGGCGAGGAAAGTCTCTCCCGAGCAAGCTGGAAGCCCTACCGAAGGGAAGGCG 2453
Qy 2259 GTCTCCACTTTGGAGTCAATTTAAAGATTAAGTCACTCCAGAAAGAAATCCAGTCAAAA 2318
Db 2454 GTTTCCACTTTGGAGTCAATTTAAAGGTTAGTCAAGCCCAAGAAAGAAATCAAGTCCAAG 2513
Qy 2319 CTGGAAGAGAAAGCCGAAGAC-----TCTAGTGTAGAGAGGTTGTCACCTGAG 2366
Db 2514 CTGGAAGAGAAAGCGAAGACTCCATAGCTGGGTCTGTGTAGAACATTCCACTCCAGAC 2573
Qy 2367 ATCGAACCGAGTAGAGAAAGATCTTGGGTTTCCATTAAAGAAATTCATCCCCGAGCGCG 2426
Db 2574 ACTGAAACCCGTTAAAGAAAGATCTCTGGGTCTCAATCAAGAAAGTTTATTTCTTGGACGAGG 2633
Qy 2427 AAGAAAGGCGAGACGCGGAAGAGCAAGACCACTGTGGAAGCTCAGGGCCAGTGGAG 2486
Db 2634 AAGAAAGGCGCAGATGGGAAACAGAAACAGAGCCCTCTGTTGAAGACGAGGGGCCAACAGGG 2693

QY 2487 ATAAATGAGACGACCTTAATGTCCAGCCGTCGTCTCTGTGAGTATTAATGCACTG 2546
 Db 2694 GCCAACGAAGATGACTGTGATGTCCCGGCGGTGTCTCTGTGATGATGATGTCTGTA 2753
 QY 2547 GAGAGGAGAAAGATG-----AAGCCAGGGGAAATAGGGAGCTGCCAGCTGTG 2597
 Db 2754 GAAAGGGAGAAATGAGAGCACAGCAAGCCAAAAAGGCGCAGACAGCCCCGAGCAAG 2813
 QY 2598 GGGGCTGTGATCTGTTCGAGAGCTCAGTAAGACTGTGTCCAACCTGTAGTGTGCA 2657
 Db 2814 GCACCCACTGAGGTGTCCAGAGAGCTCAGCGAGTCAAGTTTCAATGATGACAGACT 2873
 QY 2658 GTCAATTGATGGACCAAGGCGACTCACTGTCCAAAGCCGTCCTCTTGGATATCC 2717
 Db 2874 GTCCCTGACGGGACGAGGGCAGCTACCATTTTGAAGAAAGGTCTCTCTTGGATATCT 2933
 QY 2718 GCTTCCTGAAGAACCTCTTGAACACACAGCGGGAAAGCATGCGACCTGTGAAGAG 2777
 Db 2934 GCTTCAGTGAAGAACCTCTTGAACAAAGTGAAGCTGAAGCGCACCTTTAACTGAGAG 2993
 QY 2778 GTCACTGAAAAAGACATCATTTGC---AGAAAGAACTCCTGTGTCTCACCCAGACGTTACCA 2834
 Db 2994 GTATTGGAAAAGATGATTTGCAAGAGAAACCCCCAGGTTACTGAACCTGTGCA 3053
 QY 2835 GAGGTTAAAGATGCCATGACGACATGTGTCACTCACTGAAGTGAATTTCACTCAGAACT 2894
 Db 3054 GAGAAACAGAGAGCGCCGGGGGACACGGTCTGTAGTGAAGCGGAATTTACCCCCCAACT 3113
 QY 2895 GTGACAGCAACAGACCTCAGAGGCTCTCCGTACTGAAGAAATTACGAGACATCGGG 2954
 Db 3114 GTGACAGTGAAGAACTGACAGGGCATTTGGGTTCGAAGAGAAACGGAACACTTGTCT 3173
 QY 2955 GCCCAAGAGACCAACAGACATGATGTGCCGAGTTTCCAGCTGACTGACTGCCAGACAC 3014
 Db 3174 GCTGAAGAGACCAAGAAATGTTGTCAAGATCTCCAGTTTAACTCCACACAC 3233
 QY 3015 ACAAGAGAACCCACCTTCAAGAGTGAAGATGTGTCTGATACAGAAAGAG 3074
 Db 3234 ACAGAGAGGCGCACTCCGGTGCAGAGGTGGAAGGTGCGTACTGACATAGAAAGCAA 3293
 QY 3075 GAGCGCCAGACGACAGGCACTCTCAAGCCGTTGACAAAGGTAAAGAGAGTCCGAG 3134
 Db 3294 GAGAGGCGGACTCAAGAGTCTCCAGGCAAGGCAAGAAAGTGAAGAGAAATCCAG 3353
 QY 3135 GTGCTGTCAACCC---CAGACTGTGCAGAGAAAGGGGTCAAAAGCACTGAGAAAGTTGAG 3191
 Db 3354 CTGCTGTGCACCGGTGGGCGAGAAATGTGCTTCAAGCCGTGCAAGAGCAGAGCAGAA 3413
 QY 3192 GAGGTAGAGAGACTCCGAAGTGTGCTGCTTGCAGAAAGAGAGAGAGTTATGCCGAA 3251
 Db 3414 AGACAGAGAGAGCAGGCTGAAGCGTGGGCTGGAAGAAAGAACCGAATGTATGTTGAAA 3473
 QY 3252 GAGCCCGTGCAGAGAGCTGAGACTGAGCATCTTGCACAGGGCTGTGAAGCTGGAAGGCT 3311
 Db 3474 GTAGATGCTCAGAGAGCAAAATCTAGCCTTTTACAAAGGAGTGTGAGGAGAC 3533
 QY 3312 ACTCCAGAGAGCTTG---AAGTTCTGAAGTCAAGGACAGATGTAGAC-----CAT 3359
 Db 3534 ACCCAGAAAGCTTTGAAAAAGCTCTCAAGTCAAGAGAGATGAGAGTCCAGTAGCTT 3593
 QY 3360 GTCCGCAAGTGTCA-----GTTTATCAAGCTTCAGACGCTGATGANA 3401
 Db 3594 GTAAACCATTTGTCAAGCCGAAACTTAGCTGGGTTAAATCAAGAGAGTGTGATGANA 3653
 QY 3402 CAGGCGGTGGCCCTGAGTCAATCCGAAACTTTGACAGACATGAGAACAAATGAGAGCT 3461
 Db 3654 CAGGCTATCCCTGACTCGGTGAGAAACCTCTACAGACAGTGAACCTGATGAGAGAC 3713
 QY 3462 CCCTTAGAGATTTCAGACCTGACATGGGACACAGCAAGATGAAACCATTTGACAGCCAG 3521
 Db 3714 CCGGTAGCGACTTTGACGACACAGGACACACCCAGAAAGAGAGATTTGTGAAATCCAT 3773

QY 3522 GACAGTAAAGCACTGAGCGTGTGAGGAGTCAAGGTCAAGAAAGAGAGGCGGCTACT 3581
 Db 3774 GAGAGAAATGAGTGTGATTTGTAATCCAGTCAAGGGGCAACAGAGCAGAGGCACTTCT 3833
 QY 3582 GCTCAGAAAGAGAGGCTTGCACACTAATTAATGTTCCAGGCCAGAGAAACATGAG 3641
 Db 3834 GCACAGAAAGAGAGGCTTGCAGACCTTCCAGATTTTGTGTTCCAGAAAGAACTAAAGAA 3893
 QY 3642 GAAGAACAGAGAGAGATGTTCTTGAACCTTACACAGCAAGACTTACTGTCCAGCCGTG 3701
 Db 3894 CAATCAAAAGATGAAAGACCTTGAACATACAGATTAAGAGGGTGTCAAGTGAAGCTGA 3953
 QY 3702 CCGTGTGGCAAAAGACTGAGGTGGGTCAAGGGGTGA-----GTTTGACTGGTTG 3752
 Db 3954 TCCATTTGTTCAAAGACTGAGGGGACTCAAGAGCTTACAGATGTGCTGATGAAAAAC 4013
 QY 3753 GATGAGAAAAAGTCAAAAGAAACAGAGAGTGTGTAACACTGTG-----ACCCAC 3806
 Db 4014 AAAGACGTACATTTTGAAGGACTTGAAGGGTCTATAGACACAGGCATTAACAGTCAGT 4073
 QY 3807 AGTCAAAAGGCTGTGATGTGACATATGACATGTAAGTATGAGAGTGGCCGGGTCTGAC 3866
 Db 4074 CGGAAAAAGTCACTGAAGTTGCCCTTAAAGGTGAAGGACAGAAAGAGCTGAATGTAAA 4133
 QY 3867 GAAAGGAGGTACTGAAGTGCAGAG-----TCTTACGCTGGAAGAGGGA 3911
 Db 4134 AAGATGATGTCTTGAACCTGACAGATCAGCTTAAGTCTCTTCAATCCCGTGAAGGA 4193
 QY 3912 GAGATGAAACTGACTGTTGAAAGAGAAAGGAGACAAAGCCAGAGCAAGTGAAGTA 3971
 Db 4194 GAGATGTGATTTCAAGTGAAGGAGGAGAAACAGAACAGAGCCCAATGTGAATGAA 4253
 QY 3972 GAAG---TGACGCGAAACAGCGCTCTCTGAGCATGAAGAACCTTACGGAGCCAGTC 4028
 Db 4254 GAGAAAGCTTGAAGCAGAAACGCTGTTAACGTATCTGAAGGTCAGTAAGCACTCTCTC 4313
 QY 4029 CTGACCTTGACATGCCAGCTCAGAGAGGGGAGAGGACTGGAAGCCTTGGAGGAAGC 4088
 Db 4314 CAGACATGTAATGTGCCATATATGATGGGCAAGAAAGTCAAGTCTTGGAAAGAGAC 4373
 QY 4089 CTTTCTC---TCCAGACCAAGCAAAAGCAAGTGTGACTAGAGTTCAAGTTCAAAAGCTG 4145
 Db 4374 CTTCTCCCTGCTTAAGTCAAGAGAGGAGGATGACACAAATTTCAAGTTCAAGGCTCT 4433
 QY 4146 GACACAAAGTCACTCAACACGACGAAGCTGTGAAAA-----GTCATAGAAACGGTT 4199
 Db 4434 GAGGATATTACTCTTAACGCGCTGCAGAGAGGAAAGTCTTAAGGAAATGTGCC 4493
 QY 4280 GTGATTTCAAGAGCAGGTGAAGTCCAGAGTGTGAGGTGACACTTAATTAACAGCTGAG 4259
 Db 4494 AACATTTTGAAGACAGGTGAAGCTTGAAGCTTCAAGGTGACATTTTATGTTCTGGAAGAG 4553
 QY 4260 AAGTCTTGTCAACGGGTGGCACTGTGACTCTTCAAGATGACAGACACCGTACCCCTG 4319
 Db 4554 AATCTCTGAAAAAATGAAGACTTTGGCCGCTCATCAAGGGAAGATGCTGTGCCACA 4613
 QY 4320 GGGCTGAGTCTCAGGAGAAATCCATCCATATATGTAATCTCTGCTCTGAAAGCAC 4379
 Db 4614 GGGCCCGACTGTCAAGGCAAAATTCACACAGTGAATGATATGTCTTACTTCAAGAAAGGC 4673
 QY 4380 CTACATCTGTACTCAAGAGAAATTAAGGCATCCAGAGAGAGCGATTCAGAGGAAG 4439
 Db 4674 TTAAGTTCGACCTTGAAGGAGAGAAACCAATCACTGAAGTGAAGTCAAGATGAATGC 4733
 QY 4440 GACAAAGCAGATGTGTGCTGATGCTGACCGGCAAGAGAGATTAACGAATGAAAAATGC 4499
 Db 4734 GATGAGCAGGTGTGCTGTGC-----AGAGGTCAAAAGTGAATGAGCAATTTG---AGGAT 4784
 QY 4500 CTCAAGGCTGAACCTGAGATCTTGAACCTTGAAGTAAAGCAACAGATTTGTCTGAAC 4559
 Db 4785 TTAGAGCTGAAGAAATGGGATTTTGAATCTTGAAGACCAAAACAGTAAACTTGTCCAAAC 4844
 QY 4560 GTCATTCAGACAGCGCTTGAACAGTTCGACGTAC---AGAAACAGCCCGGAAACTCAT 4616

; SEQ ID NO 2
; LENGTH: 5346
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-738-877-2

Query Match 31.6%; Score 1945.6; DB 10; Length 5346;
Best Local Similarity 67.5%; Pred. No. 0;
Matches 3253; Conservative 0; Mismatches 1324; Indels 243; Gaps 27;

QY 33 ATGGGCGGACGATGTTCCAGCAGACGAGCGAGCCCGGACGACCGGCGC---GGGAGCGAC 89
Db 1 ATGGGCGCGGAGCTCCACGAGCAGCGACGCCGAGACGCCCGCCGAGGGAGCTCC 60
QY 90 ACGCGGACGAGCTGCTGCTCAGTGGCCATGGGCCCGGACGCTGAAAGCTCGGGA----- 143
Db 61 ACGCGGCTGAGCGCCGAGCCGAGCGGCGCGGCCCTCGGCGGAGCGCGCCAGACACC 120
QY 144 ---GCAGCTGAGACCCCGCGGACCGCGGACCCCGCCACCAAGCTCCACAGAAATGGC 200
Db 121 ACCGCGGACCCCGCATCGCTGCTCGGACCCCGCCACCAAGCTCTACAGAAATGGT 180
QY 201 CAGCTGCTTCTGTCACGCGGCTAGCTGAAACAAGAGATGTCTCATGTCCAAAGAGAAAAC 260
Db 181 CAGCTGTCCACCATCATATGGCTTAGCTAGCAAGATGAGCTCAAGCTCCAGAGGGGTAC 240
QY 261 CAGAGGG-----GCAGAGGAA 278
Db 241 CTAATATGCCAGAAAGAGACCCCTGAACGGTCAAGAGCCCTTAAACAGCCAGAGAGAA 300
QY 279 GAAGTCCTGTGATGAGGATGTTGGAACAAGAGATGTAAGAGAAAACAACCA 338
Db 301 GAAGCATTTGTCAACGAGGTTGGAACAAGAGATGTAAGAGAAAACAACCTCC 360
QY 339 GTTGAAGAAATGGCGGCAACTCCAAGCTGTTGAAGATATCAAAAGATGGGCGAGAG 398
Db 361 GATTAAGAGATGGCTACTTAAGTACGGGTTTTCAGACATACAGATGATGGGCGAGAG 420
QY 399 GAGACATCAGAAATATTAATGAACAGATCCCTGCTTCAGAAAACAATGTGGAAGAAATG 458
Db 421 GAGA---ACCGAAATATGAACAGATTCCTTCTTCAGAAAACAATTTGAAGAGCTAAC 477
QY 459 CAGCTGCTGAGTCCCGAGCTATGATGTTGGCTTCAGAAAAGTATTAAATTTGTTGGT 518
Db 478 CAACCCACTGAGTCCCGAGCTATGATGTTGAATTTAAGAGTGTAAAGTTTGTGGC 537
QY 519 TTTAAATTCACGCTGAAGAGATTAATAATGAAGTCAAGTACTGCTCAACTACTCAT 578
Db 538 TTTAAATTCAGTGAAGAAAGATTAAGACAGAAAGCTTCAGACTGTTCAGCTTACTCAT 597
QY 579 GTCAAGAGATGAAGGCGAGAGGCGAGAAAGCTCTGTTCGAGCTGAGACCAACAGAGAG 638
Db 598 GTGAAGAAAGATGAAGGAGGAGGAGAGCA-----GGGGCTGGCGAACCAACAGAGAG 648
QY 639 CCCAGTGTGAGAGCTCCGTCGGAAGTTCAGATCCAAAGAAATGAGCTGAAGCAATCC 698
Db 649 CCCAGCTTGGGGCT-----GGAAGAGCAGCATCCAAAGAAAGCGAAACCAAAACATCT 702
QY 699 ACAGAGAGCAAGAGGACCCCTGAAGCAAGAAACAGAGACACAGAAATCCCCCTTCA 758
Db 703 ACAGAGAAACCCGAAAGAGACCTTGAAGCTTGAAGCAAGCCAGCAAAATTTCTCCCA 762
QY 759 GCGGATCTGATCAAGCGGCTGAGAGAGAGCCAAAGATGAAGAGAGAGAAACAAAGAG 818
Db 763 GCGGATCTGGGCAA---GCAGTGAGAGATGCAAAAGAGAGAGAGAGAAACAAAGAA 819
QY 819 AAAGAGCCCAAGTCCCCAGAAATCCCGAGACGCCAGTCAACAGTGAAGCAATCT 878
Db 820 AAAGAACTTGAAGAGTGTGCAAAATCTCCGACTAGTCCGCTGACCGTGAACCAAGATTA 879
QY 879 TCCCTTCAGAAATTTCTTCACTCAGCGTTGGGCGGCTGGCGGAAGAGCAAGCTTCAAG 938
Db 880 ACCTTCAAAAAATTTCTTCACTCAGAGTTGGGCGGCTGGCGGAAGAGCAAGCTTCAAG 939

QY 939 AAATCAAAAGAGATGATCTTGGAAACTCCCGAGAAAGAAAGAGCAAGAGCCAGAAA 998
Db 940 AAGCGAAGAGAGATGAAGTGAAGCTTCAAGAAAGAAAGAAAGAAAGCAAGAGCCAGAAA 999
QY 999 GTAGACGAGAGAAAGAAAGAAAGACAGAGCGAGCTCCGAGGA----- 1042
Db 1000 GTAGACACAGAGAGAGACGGAAGAGCAGAGGTTCTCTCGAGAAACTGACCGCTCCGAG 1059
QY 1043 -----GCAGAGCCCGCAGAAAGACACAGACCAAGGCGAGTGTGACAGACTAC 1091
Db 1060 CAAGCCACCCACAGAGAGCGCGGAGAAAGTGGCCACAGAGCCCGGTTATCAGCTGAATAT 1119
QY 1092 GAGAAAGTGAAGTCTCTTGGAAAGACCAAGTGTGACCTCGAGGCAATGTCAAGAG 1151
Db 1120 GAGAAAGTGAAGTCTCTTGGAAAGACCAAGTGTGACCTCGAGGCAATGTCAAGAG 1179
QY 1152 AAGTGTCTCTTGGAAAGAGTGTGATGAGAAAGTGAAGGCCCAAGAA--- 1208
Db 1180 AAACCTGCTCGTTGGGACAGAAAGTGTGATGAGAAATGAAAGTCCACCAAGAGAG 1239
QY 1209 GTTGTGAGAGGTCACAGTGAACCGTGAAGAGACAGAGAGAGACAGAGAGAGAGAG 1268
Db 1240 GTTGTGCGGAAGTCCAGTCAAGTCAAGCAGTGAAGAGAAACCGAAAGACAGAAAC--- 1295
QY 1269 GGAAGGCTGAAGGGGCGTGTGTGAAAGAAACAGAGAAATCTTGTCCCTGAGAA 1328
Db 1296 -----GAGGTGAAGAAACACAGAGGCTGTGTCCAGCTGAAGAA 1335
QY 1329 CTGGCTGAGCCCGAGAGGTCCTCCAGAAAGCTGAGCTGTGAGAGCTGATGAAGAGC 1388
Db 1336 TTGTTGAATGATTCAGAACTTCAGAAACCGAACCTTGCAGAGAGCTGTGAAGTCTC 1395
QY 1389 AGAGAGATGTGTCTCTGAGAGACACACTCACTGACAGACTTAAGTCTGAAGAG 1448
Db 1396 AAAGAAAGTGTGTTCGAGAGAGACCTTACACAGGAGTGAACCTGATCTGATGAG 1455
QY 1449 AAGAGCTGCCCAACACCCAGAAAGCATTCAGTGAAGTGAAGTGTCTCTCTCAG 1508
Db 1456 AAGGTGTGTCAAAACCCCGAAAGGCTGTGTGAAGTGAAGTGTGTCTGATCAAG 1515
QY 1509 GAAAGATCAAGTATCAGGAGAGTCCCTTGAAGAACTCTTCACTAGTCTCAGGCTTAAAG 1568
Db 1516 GAGAGAAATGAAGGTGACAGAGAGTCCACTTAAGAACTTTTACAGACTGGCTTAAAG 1575
QY 1569 AAGCTGTCTGGAAGAGACAGAAAGGAGAAACGAGAGGTGGGAGAGAGAGACTTGA 1628
Db 1576 AAGCTTCTGGAAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1632
QY 1629 GAATTCACACATTCACACCGAATCCCGAGAGTGTGATGAGCAAGAGAGAGAGAGC 1688
Db 1633 GAGCACTCAGTGTCCAGCCGATTTCTCGGACAGCCAGAGAGAGAGAGAGAGAGAGC 1692
QY 1689 TCTGCTGTCTCCCGAGAGAGCTGAGAGAGCAAGTGTCTGAGAAAGGCGCGCTGAA 1748
Db 1693 TCTGCTGTATCCCTGAGAGAGCCGAGAGATCAAGTGTCTGAGAAAGGCGCTTGAAGAG 1752
QY 1749 GCACCCGAGATGGGAGAGCTGAGAGAGAACTACTTCCGATGAGAGAGAGAGAGAGAA 1808
Db 1753 GTGACAGAGAGATGGGAGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1812
QY 1809 GGGATCACTCCCTGGGCACTCTTCAAAAGATGTGACACCCAGAGAGAGAGAGAGAG 1868
Db 1813 GGTGTCACTCCCTGGGCACTTCAAAAGATGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1872
QY 1869 CTTCTGAGAGTGAAGAGAGAGAGTGAAGAGAGTCAAGAGAGAGAGAGAGAGAGAGAGAG 1928
Db 1873 CTTTGGAGAGAGATGAAGAGAGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1932
QY 1929 ACTGTAGACAGATGTCAAGAAATCAGAGATGAGTCAAACTGTTGTGAGAGAGAGAG 1988
Db 1933 ACCGAGACACAGAGCTCTGAATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1992

QY 1989 CCAGGAGNACAAAGCGTAGGTGGATACCTCAGTGTCTGGGAAGCACTGATTTGTGTC 2048
Db 1993 CCGGAAGAACAAAGCGCAAGGTGGATACCTCAGTATCTTGGGAAGCTTTAAITTTGTGTG 2052
QY 2049 GGATCATCCAAAGAGAGCAAGGAAGGCATCTCTTCCAGATGATGAAGGAGGCCCAAGG 2108
Db 2053 GGATCATCCAAAGAAAGCAAGGAAGGTCTCTTCTGATGAGGAAGGGGACCAAAA 2112
QY 2109 AACTGGGAGGGGACAGTCTACAGAGCAAGGAGGCCCAAGCAAAAGAACCGGAACA 2168
Db 2113 GCAATGGGAGGAGACCAACAGAAAGCTGATGAGGCCGGAAGAACAAAGAGAGCGGGACA 2172
QY 2169 GACGCTCTTCTGCGCAGCACCCAGGAGCAGACCAAGGCCAGGAAGTCTCTCACCCGAG 2228
Db 2173 GACGGGATCTCTTGTGTCTCCAAAGACATGATCCAGGGCAGGAAGTCTCTCCCGGAG 2232
QY 2229 CCAGCGGAAGCCCTCCGAAGGGGAAGGTGTCTCCACTTGGGAGTCAITTTAAAGATT 2288
Db 2233 CAGCTGGAGCCCTACCGAAGGGGAGGGGTCTTCCACTGGGAGTCAITTTAAAGATT 2292
QY 2289 GTCACTCCAAAGAAAAATCCAACTCAAACTGGAAGAGAAAGCCGAAGAC----- 2339
Db 2293 GTCAAGCCCAAGAAAAATCAAAGTCCAAAGCTGGAAGAGAAAGCGAAGACTCCATAGCT 2352
QY 2340 ---TCTAGTGTAGAGCAGTTGTCCACTGAGATCGAACCGAGTAGAGAAATCTTGGTT 2396
Db 2353 GGGTCTGGTGTAGAACATTCACCTCCAGACACTGAACCCCGGTAAAGAAATCTCTGGGTC 2412
QY 2397 TCCATTAAAGAAATTCATCCCGGAGCGCGGAAGAAAGGGCAGACGGGAGCAAGAACAA 2456
Db 2413 TCAATCAAGAAAGTTTATCTCTGGACAGAGGAGAAAGGCCAGATGGGAACAAGAACAA 2472
QY 2457 GCCACTGTGGAAGACTCAGGGCCAGTGGAGATAAATGAGGACGACCCCTAATGTCCAGCC 2516
Db 2473 GCCCTGTGGAAGAGCAGGGCCCAACAGGGCCCAACGAAGTACTCTGATGTCCCGGCC 2532
QY 2517 GTGTGCTCTGTCTGAGTATATGATGCAAGTGGAGAGGAGGAAGATGG-----AAGCC 2567
Db 2533 GTGTGCTCTGTCTGAGTATGATGCTGTGAAAGGGAGAAATGGAGGCAACAGAACCC 2592
QY 2568 CAGGGGAATACGGAGTGTCCCGAGCTGTGGGGCTGTGTACGTGTCCGAGGACTCAGT 2627
Db 2593 CAAAAGCGCAGAGCAGCCCGCAGAGGACGCCACTGAGGTGTCCAAAGAGCTCAGC 2652
QY 2628 AAGACTGTGTCCACTGTGAGTGTGCGAGTCAATGATGGGACAGGGCAGTCAACAGT 2687
Db 2653 GAGAGTCAGGTTCATATGATGGCAGCAGCTGTCTGACGGGACAGGGCAGCTACCAT 2712
QY 2688 GTCGAAGAGCGGTCTCCTTGTGTGATATCCGCTTCCGTTAACAGAACCTTTGAAACACA 2747
Db 2713 ATTGAAGAAAGGTCTCCTTCTTGTGATATCTGCTTCAGTGACAGAACCTCTTTGAACAAGTA 2772
QY 2748 GCGGGAGAACGATGCGACCTGTTGAGAGGTCACTGAAAGACATCATTTGC---AGAA 2804
Db 2773 GAAGCTGAAGCCGCACTGTGTTAACTGAGAGGATTTGAAAGAGAAAGTAATTTGCAGAGAA 2832
QY 2805 GAAACTCTGTGTCAAGGACGCTTACAGAGGGTTAAAGATGCCATGACGACATGGTC 2864
Db 2833 GAACCCCGCCGTTACTGAACCTCTGCCAGAGAAACAGAGAGGCCCGGGGGCAGACGGTC 2892
QY 2865 ACCAGTGAAGTGGATTTCACTCAGAGAGCTGTGACGCCACAGAGACCTCAGAGGCTCTC 2924
Db 2893 GTTAGTGGGCGGAATTTGACCCCGAAGCTGTGACAGCTGACAGAACTGACAGGSCCATTTG 2952
QY 2925 CGTACTGAAGAGTTACCGAAGCATCGGGGCCGAAAGACCAAGACCATGATGGTGTCCGCA 2984
Db 2953 GGTTCGAAGAGGAACCGAAGCATCTGCTGTGAAGAGACCAAGAGAAATGGTGTACGCA 3012
QY 2985 GTTTCCAGCTGACTCTCCAGACACCAAGAGGAAGCCACCCAGTTTCAGGAGGTA 3044
Db 3013 GTCTCCAGTTAAACGACTCCCGACACCAACAGAGAGGGCCACTCTCGGTGCGAGAGGTG 3072
QY 3045 GAGAGTGGTGTGTAGATACAGAAAGAGAGGAGCGCCAGACGCGAGGCCATCTCTCAAAGCC 3104

Db 3073 GAAGGTGGCGCTACTGACATAGAAAGCAAGAGAGCGGACTCAAGAGGTCTCTCCAGGCA 3132
QY 3105 GTTTCAGACAAAGGTGAAGAGGAGTCCAGGTGCTCTGCAACC---CAGACTGTGCAGAGA 3161
Db 3133 GTGCAGAAAAAGTGAAGAGGAAATCCAGCTGCTCTGGCACCGGTGGCCAGAAAGTGTG 3192
QY 3162 ACGGGTCAAAAGACATGGAGAAAGTTGAGAGGTAGAGGAGACTCCGAAGTGTCTGGCT 3221
Db 3193 CTTTCAGCTGTGCAGAGCAGAGGCGAGAAAGACCAGAGAGCAGGCTGAAGGTCGGGT 3252
QY 3222 TCCGGAAGAGAGGAGACGTTATGCCGAAGAGACCGGTGCAGNACTGCAGCTGAGCAT 3281
Db 3253 CTGAAGAAAGAGACGAGTGTGTTGAAAGTAGTGTCTCAGGAGGCAAAACCTGAGCCT 3312
QY 3282 CTTTCAGCAGGGCTCTGAGACTGGACAGCTCTCCAGAGAGCCTTG---AAGTTCTCTGAA 3338
Db 3313 TTTTACACAGGGAAGGTGGTGGGCGACACCCCGAGAAAGCTTTGAAAAGCTCTCTCAA 3372
QY 3339 GTCAAGCAGATGTA-----GACCATGTGCGCCACGTGCCA----- 3373
Db 3373 GTCAAGAGAGCATAGAGTCCAGTGTGTAACCACTTGTCAAAGCCGAAACCTTAGCT 3432
QY 3374 --GGTTATCAAGCTCCAGCAGCTGATGGAACAGGCCGCTGCGCTTGTAGTCTATCGAAGACC 3431
Db 3433 GGGGTAAAAATPCACAGGAGATGGTGTGGAACAGGCTATCCCGCTGACTCGGTGGAAACC 3492
QY 3432 TTGACAGACAGTGCAGACAAATGGAAGCACTCCCTTAGCAGATTCAGACACTGCAGATGGG 3491
Db 3493 CTTACAGACAGTGCAGACTGATGGAACACCCCGTAGCCGACTTTGACGCCACCGGACACA 3552
QY 3492 ACACAGCAAGATGAAACCATTTGACAGCAGGACAGTAAGACCACTGCAGCTGTCCAGGACG 3551
Db 3553 ACCAGAAAGACGAGATTGTGGAATCCATGAGGAGAAATGAGGTGCGCATCTGGTACCCAG 3612
QY 3552 TCACAGTGCACAGAAAGAGGCGGCTACTGTCTCAGAAAGAGAGCTTCGACACTACCT 3611
Db 3613 TCAGGGGGCACAAGAGCAGAGGAGTTCCTGTCACAGAAAGAGAGGCTTCGACGACCTTCC 3672
QY 3612 AATAATGTTCCAGCCCAAGGAAGAACATGGGGAAGAACCAAGGAAGAGATGTTCTTGAACCT 3671
Db 3673 AGTTTGTGTTCAGGAGAAACTAAGAAACAATCAAGATGGAAGACACTTAGAGCAT 3732
QY 3672 ACACAGCAAGAGCTTACTGTGCGAGCGCTGCCGCTTCTGGCAAGAGCTGAGGTGGGTCAA 3731
Db 3733 ACAGATAAAGAGGTGTCAAGTGGAAACTGTATCCATTCTGTCAAAGACTGAGGGGACTCAA 3792
QY 3732 GAGGGTGA-----GGTTGACTGGTGGATGGAGAAAGTCAAAGAAACAGGAG 3782
Db 3793 GAGGCTGACCAAGTATGCTGATGAGAAACCAAGACGTAACCATTTTTCGAAGGACTTGAG 3852
QY 3783 GTGTTGTACACTCTGG-----ACCCACAGTCAAAGGCTGCTGATGTGACATATGAC 3836
Db 3853 GGGTCTATAGACACAGCATACAGTCAAGTCCGGAAGAGTCACTGAAGTTGCCCTTAAA 3912
QY 3837 AGTGAAGTCAATGGGAGTGGCGGCTGTTCAGAAAGAGAGTACTGAAAGTCAGAG----- 3892
Db 3913 GGTGAAGGACAGAAAGCTGATGTAAAAGGATGATGCTCTTGAACCTGCAGAGTCAAC 3972
QY 3893 -----TCTTAGCTTGGAGGGGAGAGATGGAATCTGACGTTGAAAAAGGAGAAA 3941
Db 3973 GCTAAGTCTCTCCATCCATCCCGTGGAGAGAGATGGTAGTTCAAGTTCGAAGAGGAGAAA 4032
QY 3942 AGGAGACAAAGCCAGAGCAAGTGAAGTGAAG---AAGGTGAGGAGGAAACAGCCGCTCCT 3998
Db 4033 ACAGAAAGCAGAGCAACCCCATGTGAAATGAAGAGAACTTGAGACGAAACAGCTGTTACC 4092
QY 3999 GAGCATGAAGAAACCTACGGGAAGCCAGTCTGACACTTTCACATGCCAGCTCAGAGAGG 4058
Db 4093 GTATCTGAAGAGGTCAAGTGAAGCAGCTCTCCAGACAGTGAATGTGCCCATCATAGATGG 4152
QY 4059 GGAAGGCACTGGGAAGCCCTTGGAGGAAGCCCTTCTC---TCCAGACCAAGACAAAGCA 4115


```

Db 4153 GCAAGAGAGTCAGCAGTTGGAGAGAGCCCTCTCTCTGCTAGTCAAGAGAGGCA 4212
Qy 4116 GGTTCATAGAGTTCAAGTTCAAGCCTGAGACACAGTCACTCAAAACGAGAGCT 4175
Db 4213 GTATGCAACAAATTCAGATTCAAGCTCTGAGCATCATCTCAACGCGCTGCA 4272
Qy 4176 GTGAGAAAGTCAT-----AGAAACGTTGATTTCAAGACAGGTGAAGTCCAGAG 4229
Db 4273 GAGGAGGAGAAAGTCTTGAAGGAGAACTCCCAACATTTTGAAGAGGAGAGGAG 4332
Qy 4270 TGTGAGGTGACACTTATTAACAGCTGAGAAAGTCTCTGCAACGGGTGAGCAGTGA 4289
Db 4333 CTTGAGGTGACACTTATTAAGTTTGAAGAGAAATCTCTGAAAAAATGAAGATTTGCC 4392
Qy 4290 CTTGAGGTGACAGAGACAGTACCCCTGGGGCTGAGTCTGAGGAGAGATCCCA 4349
Db 4333 GCTCATCCAGGGAGAGATGCTGTGCCACAGGGCCGACTGCAAGCAAAATCGACACA 4452
Qy 4350 ATCATAGTAACTCTGCTCTCTGAAAGCACCCTCAATCTGACCTCAAGAGAAATTA 4409
Db 4453 GTGATAGTATCTGCTAATCAAGAAAGGCTTAAGTTCGACCTGGAAGAGAGAAAC 4512
Qy 4410 GCATCCAGAGAGAGAGATCAAGAGAGAGCAAGCCAGATGCTGCTGATGCTGAC 4469
Db 4513 ACATACCTGAAGTGAAGTCAGATGAATGATGATGAGCAGGTTGCTTGC-----AGAG 4566
Qy 4470 GGCAAGAGAGATGACAGCAATCGAAAAAGTCTCAAGGCTGAACCTGAGATCTGAACTT 4529
Db 4567 GTCAAGAGAGATGATGAGCAATG---AGATTTAGAGCCTGAAATGAGGATTTTGAAGCTT 4623
Qy 4530 GAGAGTAAAGCAACAGATTGTGCTGAAGCTCAATTCAGACGCGCTTACCAAGTTGCA 4589
Db 4624 GAGACCAAAAGCAGTAACTTGTCCAAAACATCATCCACACAGCGTTGACCAAGTTGTA 4683
Qy 4590 CGTAC---AGAAACAGCCCGGAACTCATGCTTATGATTCACAGCCAGGTTCTGCA 4646
Db 4684 CGTACGAGAGAAAGCAGCCACCCAAATGTTGACGTCTGATTTACAGACACAGCTACG 4743
Qy 4647 TGCAGGCTTGACAGCAGGGA 4666
Db 4744 ATAAAGCTGACAGCCAGGA 4763

```

RESULT 6
US-09-917-800A-1223/C

; Sequence 1223, Application US/09917800A
; Patent No. US20020119462A1

; GENERAL INFORMATION:

; APPLICANT: Mendrick, Donna

; APPLICANT: Porter, Mark

; APPLICANT: Johnson, Kory

; APPLICANT: Casle, Arthur

; APPLICANT: Elashoff, Michael

; APPLICANT: Gene Logic, Inc.

; TITLE OF INVENTION: Molecular Toxicology Modeling

; FILE REFERENCE: 44921-5038-US

; CURRENT APPLICATION NUMBER: US/09/917,800A

; CURRENT FILING DATE: 2001-07-31

; PRIOR APPLICATION NUMBER: US 60/222,040

; PRIOR FILING DATE: 2000-07-31

; PRIOR APPLICATION NUMBER: US 60/222,880

; PRIOR FILING DATE: 2000-11-02

; PRIOR APPLICATION NUMBER: US 60/290,029

; PRIOR FILING DATE: 2001-05-11

; PRIOR APPLICATION NUMBER: US 60/290,645

; PRIOR FILING DATE: 2001-05-15

; PRIOR APPLICATION NUMBER: US 60/292,336

; PRIOR FILING DATE: 2001-05-22

; PRIOR APPLICATION NUMBER: US 60/295,798

; PRIOR FILING DATE: 2001-06-06

; PRIOR APPLICATION NUMBER: US 60/297,457

; PRIOR FILING DATE: 2001-06-13

; PRIOR APPLICATION NUMBER: US 60/298,884

```

; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,459
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 1740
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1223
; LENGTH: 563
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 A123818
US-09-917-800A-1223

```

Query Match 8.9%; Score 550; DB 10; Length 563;

Best Local Similarity 99.1%; Pred. No. 1.9e-135; Indels 0; Gaps 0;

Matches 553; Conservative 0; Mismatches 5;

```

Qy 5601 ATAGCAGCGCTTTGAAAGCTCCAGCCTCACTGTAACCTGACGAAACAGATACATT 5660
Db 563 AAAGCAGCGCTTTGAAAGCTCCAGCCTCACTGTAACCTGACGAAACAGATACATT 504
Qy 5661 CTTGCAAGAGAGAGACAAATCTTTTAAAGTTTACTGATCTTAAATCTGAGGCTTCT 5720
Db 503 CTTGCAAGAGAGAGACAAATCTTTTAAAGTTTACTGATCTTAAATCTGAGGCTTCT 444
Qy 5721 AGTCTCTGAAAGGTGTTGTTTCTCTATGACAGCGAGCTCAGAAATPAAACCCCATTT 5780
Db 443 AGTCTCTGAAAGGTGTTGTTTCTCTATGACAGCGAGCTCAGAAATPAAACCCCATTT 384
Qy 5781 TGAACATCCAGATGATGATCCCAATATACATGATTTTTCCTTATGAGCTGCTTAAATGATGCA 5840
Db 383 TGAACATCCAGATGATGATCCCAATATACATGATTTTTCCTTATGAGCTGCTTAAATGATGCA 324
Qy 5841 TCCAGGTGAAAGAGAGTCTCTCTGTCGATGATTAAGCCCTGCTTAAATGATGCA 5900
Db 323 TCCAGGTGAAAGAGAGTCTCTCTGTCGATGATTAAGCCCTGCTTAAATGATGCA 264
Qy 5901 CAATGATGATGCTTAAAGCCATGAGATGTTCTTAAATGCAAGAGATCTGTTGATGCT 5960
Db 263 CAATGATGATGCTTAAAGCCATGAGATGTTCTTAAATGCAAGAGATCTGTTGATGCT 204
Qy 5961 TTTTGTATGATCTCTCTTATGTCGACCGAATTCATATGACATGCAAGTGAAGTCTG 6020
Db 203 TTTTGTATGATCTCTCTTATGTCGACCGAATTCATATGACATGCAAGTGAAGTCTG 144
Qy 6021 TTTTGTATGATCTCTCTTATGTCGACCGAATTCATATGACATGCAAGTGAAGTCTG 6080
Db 143 TTTTGTATGATCTCTCTTATGTCGACCGAATTCATATGACATGCAAGTGAAGTCTG 84
Qy 6081 TCTTTAAGACAAATGTCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6140
Db 83 TCTTTAAGACAAATGTCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 24
Qy 6141 AAATAACATATTTGACT 6158
Db 23 AAATAACATATTTGCTT 6

```

RESULT 7
US-09-917-800A-1475

; Sequence 1475, Application US/09917800A

; Patent No. US20020119462A1

; GENERAL INFORMATION:

; APPLICANT: Mendrick, Donna

; APPLICANT: Porter, Mark

; APPLICANT: Johnson, Kory

; APPLICANT: Casle, Arthur

; APPLICANT: Elashoff, Michael

; APPLICANT: Gene Logic, Inc.

; TITLE OF INVENTION: Molecular Toxicology Modeling

; FILE REFERENCE: 44921-5038-US

; CURRENT APPLICATION NUMBER: US/09/917,800A

; CURRENT FILING DATE: 2001-07-31


```

: PRIOR APPLICATION NUMBER: 60/277,705
: PRIORITY FILING DATE: 2001-03-21
: NUMBER OF SEQ ID NOS: 116
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO: 76
: LENGTH: 720
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of A
: OTHER INFORMATION: oligonucleotide
US-10-101-487-76

```

Query Match	1.7%;	Score 106;	DB 9;	Length 720;
Best Local Similarity	49.9%;	Pred. No. 6.5e-18;		
Matches 349;	Conservative 0;	Mismatches 345;	Indels 6;	Gaps 3.
QY 946	AAGAGGATGATCTGGAACCTCCGAGAAAGAAAGACCAAGAGCCAGACAAAAGTGTGACG	1005		
Db 710	AAGAGGAAGAGAGGAAGAGGAGGAAGAGGAAGAGAGAGGAGGAAGAGGAGGAAGAGAGG	651		
QY 1006	AGGAAGAAAAGAAAGAAGACAGAGCCAGCTCGAGAGGAGAGAGACCGGACGAAGAACACAG	1065		
Db 650	AGGAAGAGAGGAAGAAAGAAAGAGAGGAAGAGAGGAGAACAGAAAGAGAGAGAGAGAGG	591		
QY 1066	ACCAGGCCAGGTGTCAGACAGACTACGAAGAAGTGGAGCTGCCTTTGGAAAGCCAGGTGG	1125		
Db 590	AAGAGGAAGAGAGAGGAAGAGAGAGGAAGAGGAAGAGAGGAAGAGAGGAAGAGGAAGG	531		
QY 1126	GTGACCTGGAGGCATCGTCAGAGAGAAAGTGTCTCTTTGGCA--ACGGAAGTGTGGT	1182		
Db 530	AGGAAGAGGAGGAAGAGGAAGAGGAGGAAGAGGAAGAGGAAGAGGAAGAGGAAGAGG	471		
QY 1183	ATGAGATGATGGAAGCCCAAGAAAGTTGTCAGAGGTCCACGTGACACCGTGGAGA	1242		
Db 470	AAGAGGAAGAGAGGAAGAGGAAGGAAGAAAGAGAGGAAGAGAGGAAGAGGAAGAGG	411		
QY 1243	AACACAGAGAGAGAGCAGGAGAGGAGAGAGGCTGAAGGGGCGTGTGTGTGAAGAA	1302		
Db 410	AGGAAGAGAGGAAGAGGAAGAGGAAGAGGAAGAGGAAGAGGAAGAGGAAGAGGAGG	351		
QY 1303	CAGGAAGATCTTGTCCCTCTGAGAAACTGCGTAGGCCCAAGAGGTCCCCAGGAAGCTG	1362		
Db 350	AAGAGGAAGAGGAGGAAGAGGAAGGAAGGAAGGAAGAGGAAGAGGAAGAGGAAGAGG	291		
QY 1363	AGCTGTCTGAGAGCTGATGAAGACAGAGAGATGTGTCTCTCTGAGAGAGACCACTC	1422		
Db 290	AGGAAGAGAGGAAGAGGAAGAGGAAGAGGAAGGAAGGAAGAGGAAGAGGAAGAGG	231		
QY 1423	AACCTACACACTTAAGTCTCTGAAGAAACGCTGCCCAACCCAGAGAGCATTTGTCA	1482		
Db 230	AAGAGGAAGAGGAGGAAGAGGAGAGGAAGAGGAAGAGGAAGAGGAAGAGGAAGAGG	171		
QY 1483	GTGAGGTGAGATGCTGTCTCTCAGGAA-AGATCAAGGTACAGGAAAGTCTTTGAAG	1541		
Db 170	AGGAAGAGAGGAAGGAAGGAAGGAAGGAAGAGGAAGGAAGGAAGGAAGGAAGG	111		
QY 1542	AAACTCTTCAGTAGCTCAGGCTTTAAAGAAAGCTGTCTTGGAGAGACGAAGGCGGAAACG	1600		
Db 110	AAGAGGAAGAGGAAGGAAGAGGAAGGAAGGAAGAGGAAGGAAGGAAGGAAGAGG	51		
QY 1601	AGGAGGTGGGGAGACGAAGAGCCTGGAAGAAATACCAACA	1639		
Db 50	AGGAAGAGAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGG	11		

RESULT 10
US-10-101-487-71
Sequence 71, Application US/10101487
Patent No. US20020169125A1
GENERAL INFORMATION:
APPLICANT: LEUNG, DAVID W.
APPLICANT: BERGMAN, PHILIP A.

```

? APPLICANT: LOFOUJIST, ALAN
? APPLICANT: PIETZ, GREGORY E.
? APPLICANT: TOMPKINS, CHRISTOPHER K.
? APPLICANT: WAGGONER, JR., DAVID W.
? TITLE OF INVENTION: RECOMBINANT PRODUCTION OF POLYANIONIC POLYMERS AND USES
? TITLE OF INVENTION: THERDOF
? FILE REFERENCE: 077319/0329
? CURRENT APPLICATION NUMBER: US/10/101,487
? CURRENT FILING DATE: 2002-03-20
? PRIOR APPLICATION NUMBER: 60/277,705
? PRIOR FILING DATE: 2001-03-21
? NUMBER OF SEQ ID NOS: 116
? SOFTWARE: PatentIn Ver. 2.1
? SEQ ID NO 71
? LENGTH: 522
? TYPE: DNA
? ORGANISM: Artificial Sequence
? FEATURE:
? OTHER INFORMATION: Description of Artificial Sequence: Synthetic
? OTHER INFORMATION: oligonucleotide
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (1)..(522)
? US-10-101-487-71

```

	Query Match	Similarity	1.5%: Score	89.6: DB	9: Length	522:
	Best Local	Similarity	50.2%: Pred.	1.1e-13:		
	Matches	221: Conservative	0: Mismatches	219:	Indels	0: Gaps
QY	937	AGAAATCAAAGAGATGATCTTGGAATCTCCGAGAAAGAAAGGACCAAGGCGAAGAA	996			
DB	5	AGGAAGAGAGAAAGAAAGAGAGAGAAAGAAAGAAAGAGAAAGAGAAAGAGAAAG	64			
QY	997	AAGTAGACGAGAGAAAGAAAGAAAGCAAGAGCCAGCTCGAGAGAGCAGAGCCGGCAG	1056			
DB	65	AGGAAGAGAGAGAAAG	124			
QY	1057	AAGACACAGACCAGGCCAGGTTGTCAGACACTACAGAGAGTGAGCTGCCCTTGGAG	1116			
DB	125	AGGAAGAGAGAAAGAAAGAGAGAGAGAAAGAAAGAGAGAGAGAGAGAGAGAGAG	184			
QY	1117	ACCAGGTTGGATGACTCTGAGGACATGTCAGAGAGAAAGTGATGCTCTTTGGCAACGAA	1176			
DB	165	AGGAAGAGAGAGAAAG	244			
QY	1177	TGTTTGATGAGAGATGAAAGTGAAGCCCAACAAGAGTTGTTGCAGAGSTCCAAGTGAACCCG	1236			
DB	245	AGGAAGAGAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	304			
QY	1237	TGGAAAGACAGAGAGAGGACAGGAGAGGACGAGAGAGAGGCTGAAGGGGGCGTGAGGTG	1296			
DB	305	AGGAAGAGAGAGAAAG	364			
QY	1297	AAGGAACAGAGAGATCTTGCCCCCTGAGAAACTGGCTGAGGCCCAAGAGATCCCCCAG	1356			
DB	365	AGGAAGAGAGAAAG	424			
QY	1357	AAGCTGAGCGTCTGAGAG	1376			
DB	425	AGGAAGAGAGAAAGAGAGAG	444			

RESULT 11
 US-10-101-487-73/c
 ; Sequence 73, Application US/10101487
 ; Patent No. US20020169125A1
 ; GENERAL INFORMATION:
 ; APPLICANT: LEUNG, DAVID W.
 ; APPLICANT: BERGMAN, PHILIP A.
 ; APPLICANT: LOFOUISI, ALAN
 ; APPLICANT: PIETZ, GREGORY E.
 ; APPLICANT: TOMPKINS, CHRISTOPHER K.
 ; APPLICANT: WAGGONER JR., DAVID W.


```
;; TYPE: DNA
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;; OTHER INFORMATION: oligonucleotide
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(537)
US-101-487-106

Query Match          1.5%; Score 89.6; DB 9; Length 554;
Best Local Similarity 50.2%; Pred. No. 1.2e-13;
Matches 221; Conservative 0; Mismatches 219; Indels 0; Gaps 0;

QY 937 AGAAATCAAAAGAGATGATCTGAAATGCGGAGAGAGAAAGCAAGGCAAGAA 996
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 17 AGGAAGAGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 76

QY 997 AAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1056
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 77 AGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 136

QY 1057 AAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1116
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 137 AGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 196

QY 1117 ACCAGGTTGCTGACCTGAGAGGATGTCAGAGAGAGAGAGAGAGAGAGAG 1176
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 197 AGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 256

QY 1177 TGTTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1236
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 257 AGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 316

QY 1237 TCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1296
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 317 AGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 376

QY 1297 AAGGAACAGAGAGATCTTGTCCCTGAGAGAGAGAGAGAGAGAGAGAGAG 1356
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 377 AGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 436

QY 1357 AAGCTGAGCTCTGAGAGAG 1376
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 437 AGGAAGAGAGAGAGAGAGAGAG 456

RESULT 14
US-09-864-761-20733
; Sequence 20733, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmics-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
```

```
;; PRIOR APPLICATION NUMBER: PCT/US01/00667
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 20733
;; LENGTH: 575
;; TYPE: DNA
;; ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC005822.1
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 13
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 13
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 11
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 11
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 14
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 9.8
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 12
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 9.4
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 11
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 12
OTHER INFORMATION: NT HIT: APL8525.1, EVALUATE 4.80e+00
OTHER INFORMATION: EST_HUMAN HIT: BE071749.1, EVALUATE 5.80e+00
US-09-864-761-20733

Query Match          1.2%; Score 71; DB 10; Length 575;
Best Local Similarity 53.3%; Pred. No. 1e-08;
Matches 172; Conservative 0; Mismatches 150; Indels 1; Gaps 1;

QY 969 GAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1027
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 212 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 271

QY 1028 GCCAGCTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1087
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 272 GCATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 331

QY 1088 CTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1147
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 332 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 391

QY 1148 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1207
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 392 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 451

QY 1208 AGTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1267
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 452 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 511

QY 1268 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1290
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 512 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 534
```

```

RESULT 15
US-09-864-761-3972
; Sequence 3972, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864, 761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 3972
; LENGTH: 1969
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005922.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 13
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 13
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 11
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 11
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 14
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 9.8
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 12
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 9.4
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 11
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 12
US-09-864-761-3972

```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 12, 2002, 14:13:30 ; Search time 7323 Seconds
(without alignments)
13623.405 Million cell updates/sec

Title: US-09-902-432-3
Perfect score: 6160
Sequence: 1 cggccgagggagtagaag.....aaataaacatttgactac 6160

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```

EST.*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_htc.*
9: gb_est1.*
10: gb_est2.*
11: gb_htc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: gb_gss.*
18: em_gss_hum.*
19: em_gss_inv.*
20: em_gss_pln.*
21: em_gss_vrt.*
22: em_gss_fun.*
23: em_gss_man.*
24: em_gss_mus.*
25: em_gss_other.*
26: em_gss_pro.*
27: em_gss_rod.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	751.4	12.2	1153	14 D88456	D88456 D88456 Mous
2	682.8	11.1	710	14 BQ200192	BQ200192 UI-R-DQ1-
3	634.8	10.3	683	13 BI274459	BI274459 UI-R-CW0-
4	571	9.3	776	13 BG974884	BG974884 602842771
5	552.8	9.0	1269	11 BC015830	BC015830 Homo sapi
6	550.4	8.9	746	13 BM230239	BM230239 K0292D03-

C	7	550	8.9	563	9	AI233818	AI233818 EST230506
C	8	547.6	8.9	661	13	BI652823	BI652823 603298753
C	9	532.2	8.6	710	13	BM119923	BM119923 L0932E09-
C	10	529.2	8.6	668	10	BB546856	BB546856 BB546856
C	11	512	8.3	703	14	BQ176546	BQ176546 UI-M-DJ2-
C	12	501.8	8.1	610	9	AI225639	AI225639 u113909.Y
C	13	492.2	8.0	665	13	BM228830	BM228830 K0268G02-
C	14	467.8	7.6	476	9	AI103489	AI103489 EST212778
C	15	460.2	7.5	573	12	BG670384	BG670384 DRNBAE11
C	16	459.6	7.5	679	14	BQ443215	BQ443215 UI-M-EV0-
C	17	458.6	7.4	567	12	BF286252	BF286252 EST450843
C	18	456.6	7.4	623	13	BM226613	BM226613 K0230F11-
C	19	453.8	7.4	625	13	BM226974	BM226974 K0236B12-
C	20	450.6	7.3	530	9	AI156819	AI156819 u144c01.Y
C	21	442.8	7.2	601	13	BM226680	BM226680 K0231G01-
C	22	442.8	7.2	614	13	BM245338	BM245338 K0723D03-
C	23	439.6	7.1	754	13	BG966814	BG966814 602834309
C	24	437.8	7.1	745	12	BG663107	BG663107 DRA08D03
C	25	436.6	7.1	534	10	AW919250	AW919250 EST350554
C	26	434.8	7.1	593	13	BM225990	BM225990 K0220D09-
C	27	430.8	7.0	586	13	BM293486	BM293486 C0305H04-
C	28	429.4	7.0	453	10	AW535062	AW535062 UI-R-B50-
C	29	421.4	6.8	641	10	BB665546	BB665546 BB665546
C	30	418.2	6.8	577	13	BM230475	BM230475 K0295H08-
C	31	416.6	6.8	428	9	AI176113	AI176113 EST213692
C	32	415.8	6.8	574	13	BM230613	BM230613 K0298D07-
C	33	414.8	6.7	574	13	BM229218	BM229218 K0275C05-
C	34	412	6.7	589	13	BM244908	BM244908 K0717D04-
C	35	410.2	6.7	886	12	BF163549	BF163549 601769945
C	36	408.6	6.6	644	13	BI446672	BI446672 dac81f11.
C	37	406.6	6.6	674	9	AL045541	AL045541 DRP2p434K
C	38	405.8	6.6	704	13	BI645683	BI645683 603275189
C	39	400.4	6.5	562	14	BQ554344	BQ554344 H4027E12-
C	40	400.2	6.5	483	9	AA963479	AA963479 UI-R-E1-g
C	41	392.6	6.4	552	13	BM124130	BM124130 L0535F11-
C	42	385.2	6.3	599	9	AI317366	AI317366 u113909.X
C	43	384.2	6.2	550	13	BM244784	BM244784 K0715G12-
C	44	382	6.2	599	14	BQ566352	BQ566352 G158B03.Y
C	45	371.6	6.0	522	13	BM197966	BM197966 C0352A03-

ALIGNMENTS

RESULT 1
D88456
LOCUS D88456 Mouse ICR testis (J.Tsushima) Mus musculus cDNA clone A-12, 1153 bp mRNA linear EST 01-APR-1998
DEFINITION D88456 mRNA sequence.
ACCESSION D88456
VERSION D88456.1 GI:1669580
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 1153)
AUTHORS Takeo,M.M., Araki,Y., Matsunaga,A., Yokoi,A., Tsuchida,J., Nishina,Y., Noraki,M., Tanaka,H., Koga,M., Uchida,K., Matsumiya, Okuyama,A., Rochelle,J.M., Nishimune,Y., Matsui,M. and Seldin,M.
TITLE Mapping of eight testis-specific genes to mouse chromosomes
JOURNAL Genomics 46 (1), 138-142 (1997)
MEDLINE 98066773
COMMENT Contact: Junji Tsuchida
Department of Science for Laboratory Animal Experimentation
Research Institute for Microbial Diseases, Osaka University
3-1 Yamadaoka, Suita, Osaka 565, Japan
Tel: 06-879-8338
Fax: 06-879-8339
Email: tsuchida@biken.osaka-u.ac.jp.
Location/Qualifiers
1. .1153
/organism="Mus musculus"

```

/strain="ICR"
/db_xref="taxon:10090"
/clone="A-12"
/sex="male"
/tissue="testis"
BASE COUNT      367 a      285 c      299 g      201 t      1 others
ORIGIN
Query Match      12.2%; Score 751.4; DB 14; Length 1153;
Best Local Similarity 81.1%; Pred. No. 1.6e-159;
Matches 975; Conservative 0; Mismatches 157; Indels 70; Gaps 6;

OY 4081 GAGGAGGCGCTTCTCCCAAGCAAGCAAGGCTTGTGCTAAGGTTCAAGTTCAAA 4140
Db      1 GGGGAGGCGCTTCTCCCAAGCAAGCAAGGCTTGTGCTAAGGTTCAAGTTCAAA 60
OY 4141 GCGTGGACACAACTCACTCAAAAGCAAGAGCTGTGGAAGAGTTCATAGAAAGCTTG 4200
Db      61 GCTCAGACACCACTCACTCAAAAGCAAGAGCTGTGGAAGAGTTCATAGAAAGCTTG 120
OY 4201 TGATTTCAGAGCAGGTGAAGTCCAGAGTGTGAGTGCACTTATTACCAAGCTGAGA 4260
Db      121 CAATTCAGAGTGTGATGAAGTTGAGTGTGCAAGTCCCAATCATTAACCACTGAGA 180
OY 4261 AGTCTCTGCAACGGGTGGCCCTGGACTCTTCAAGATGCAAGAGACACCGTACCTGG 4320
Db      181 AGCTCTCCGAAACCGGTGGCTTACGGGACTCTTCAAGATGCAAGAGACACCGTACCG 240
OY 4321 GGCGTGAAGTCAAGCAGATTCATCCCAATCATAGTAACTCTCTCTCTGAAACACACC 4380
Db      241 NGCTGAAGTCTCAGCAGAGTCAATCCCAATCATAGTAACTCTCTCTCTGAAACATCC 300
OY 4381 TACATCTGACCTTACAGAGAAATTAAGGCATCCCAAGAGAGCGATCAAGAGAGG 4440
Db      301 TACATCTGACCTTCAAGAGAAATTAAGGCATCCCAAGAGAGCGATCAAGAGAGTGA 360
OY 4441 ACAAGCCAGATGCTGTCTCTGATGCTGAGCGCAAGAGAGTACAGCAATCCGAAAGTCC 4500
Db      361 ACAAGCCAGATGCTGTCTCTGATGCTGAGCGCAAGAGAGTACAGCAATCCGAAAGTCC 420
OY 4501 TCAAGGCTGAACCTGAGATCCCTGGAATCTGAGATTAAGAGCAAGATTTGCTGAACG 4560
Db      421 TCAGGCTGAACCTGAGATCTTGAATCTGAGATTAAGAGCAAGATTTGCTGAACG 480
OY 4561 TCATTTCAGACAGCGTTGACAGATTCGACAGTACAGAAACAGCCCGCAACTCATGCTT 4620
Db      481 TCATTCAGACAGCGTTGACAGATTCGACAGTACAGAAACAGCCCGCAACTCATGCTT 540
OY 4621 ATGATTCACAGCCAGATTCCTGCAATGCAAGCTTGACAGCGAGGAGCCCAAGATGCT 4680
Db      541 CTGATTTTCAGATTCAGATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 593
OY 4681 GGACAAAATGAAG-----ATGCCAAGTGAAGAACCCAGAGCGCGCAGC 4725
Db      594 GGACAAAATGAAGCGACCTTCAAGTCTCCCCCAAGTGAAGAACCTTGAAGCGCTAGC 653
OY 4726 CCAGAGAGGACTTGCAGTCTGACCGTTCTGAGAGCATGGGCTCAGCCTCGAAATGCT 4785
Db      654 CCAGAGAGGACTTGCAGTCTGATGATGTTCTGAAGCATGACAGCAAGGCTTCAGAAATGAT 713
OY 4786 TGCCCGGCTTGCAGTTGAAGCGCGGTGTCAAAGTAAAGCATTTGAAGAGCTCTCTCA 4845
Db      714 CACCAAGCTTGCAGTTGAAGGTGCAAGTGTCAAAGAAAGTGTGGAAGAGCTCTCTCTCA 773
OY 4846 ACCCAAGATCCAAAGAGGAGTGTGCTGATAGGCGCTCAGTCCAAAGCTTAAAGCCAGG 4905
Db      774 GTGCAAAAGT-GAAAAGGAACTGTCTGACGGCGCCCGAGACCAAAAGCTTAAGCCAGG 832
OY 4906 CAGAGGCGAGTGCCTCTGGAATCTTAACCAAGAGTCCCAAGACCAAGCGAACCAAGC 4965
Db      833 CAGAGGCGAGTGCCTCTGGAATCTTAACCAAGAGTCCCAAGACCAAGCGAACCAAGC 892

```

```

OY 4966 TAACGAGAGAGGCGATCCCCCAAAAGTTGAGTCCAGAGAGAAAGATGCTACCAAGT 5025
Db      893 TAACGAGAGAGAGGAGTGCCTTGAAGAG----- 922
OY 5026 CAGTCAAGAGAGAGAGGCGCCAGAGAGAGG---GACCTGCAAGAGCCAAAGAGAGACC 5082
Db      923 -----AATGAACAGAGGCCAGACAGAGAGAGAGAGCTTACAGAGCCAAAGAGAGACC 976
OY 5083 TGGCAGATCTTAAGATGTTAGTTGCTCATTTGTCATCTGTGAAGACCAAGATGTGAAGAC 5142
Db      977 TGACAGATCTTAAGAGTGTGTTGCTCAT-----TGTAAACTGAATGTGAAGTG 1028
OY 5143 AAGTCAGAGAGAGAGTGTGCTGTGAGCTTGAAGCCAAAGATTCAGAGCCCATGAG 5202
Db      1029 AAGTCAGAGAGAGAGTGTGCTGTGAGCTTGAAGCCAAAGATTCAGAGCCCATGAG 1068
OY 5203 ATCCAGAGAGAGAGGCGCTCCCAATGATTTCCACCAAGAGAGAGAGCCCAATTCGAG 5262
Db      1089 GTGCAAGAGAGAGAGGCGCTCCCAATGATTTCAACCTGCAAGAGAGAGCCCAATTCGAG 1148
OY 5263 GC 5264
Db      1149 GC 1150

```

RESULT 2

B0200192/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

710 bp mRNA linear EST 02-MAY-2002
 UI-R-DQ1-clu-e-23-0-UI.s1 UI-R-DQ1 Rattus norvegicus cDNA clone
 UI-R-DQ1-clu-e-23-0-UI 3', mRNA sequence.

B0200192
 B0200192.1 GI:20416657

EST.

Norway rat.

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 710)

Bonaldi,M.F., Lennon,G. and Soares,M.B.

Normalization and subtraction: two approaches to facilitate gene
 discovery

Genome Res. 6 (9), 791-806 (1996)

97044477

CONTACT: Soares, MB

Program for Rat Gene Discovery and Mapping

University of Iowa

451 Eckstein Medical Research Building Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: msoares@blue.weeg.uiowa.edu

The sequence contained an oligo-dT track that was present in the
 oligonucleotide that was used to prime the synthesis of first
 strand cDNA and therefore this may represent a bonafide poly A
 tail. The sequence tag present in the cDNA between the NotI site
 and the oligo-dT track served to verify it as a clone from the

normalized rat cell line R3327-5a library cDNA library Preparation:
 M.B. Soares Lab Clone distribution: clones will be available
 through Research Genetics (www.resgen.com)

Seq primer: M13 Forward

POLYA=yes.

Location/Qualifiers

1. 710

/organism="Rattus norvegicus"

/strain="Sprague-Dawley"

/db_xref="taxon:10116"

/clone="UI-R-DQ1-clu-e-23-0-UI"

/clone_lib="UI-R-DQ1"

/dev_stage="adult"

/lab_host="DH10B (Life Technologies)"

/note="Vector: pT7TD-Pac (Pharmacia) with a modified
 polylinker. Site 1: Not I; Site 2: Eco RI; The UI-R-DQ1
 library is a normalized Rat Cell Line R3327-5A library


```

Db 504 AAGTCTTTTAAATTAATCTATGCTTATGCTGTGGGCTTCTAGTCTCTGAAAGTGG 445
Qy 5737 TTGTTTCTATGACAGAGAGCTCAGAAATATAAAACCCATTTTGAACATCCAGATG 5796
Db 444 -TGTTTCTATGACAGAGAGCTCAGAAATATAAAACCCATTTTGAACATCCAGATG 386
Qy 5797 TCCCAATATATACATGATTTTTCCTTTTCTTTTCTTATCCAGTCCAGTTGAAAGAA 5856
Db 385 TCCCAATATATACATGATTTTTCCTTTTCTTTTCTTATCCAGTCCAGTTGAAAGAA 326
Qy 5857 GTCCTCTGTCGATGATTAAGCCCTGTCTCTTAATGATATGACAAATAGTGTGCTTA 5916
Db 325 GTCCTCTGTCGATGATTAAGCCCTGTCTCTTAATGATATGACAAATAGTGTGCTTA 266
Qy 5917 AGGCATGAGATGTTTCTTATGACAGAGAAATCTGTTGACGTTTCTTATGATGATG 5976
Db 265 AGGCATGAGATGTTTCTTATGACAGAGAAATCTGTTGACGTTTCTTATGATGATG 206
Qy 5977 TTCTATGCTGACCGAATTCATATGACATGCAAGTGAATCCTGTTCTTACAGATGTA 6036
Db 205 TTCTATGCTGACCGAATTCATATGACATGCAAGTGAATCCTGTTCTTACAGATGTA 146
Qy 6037 TTTTATGATGATCTGAGATTTGCTGTGTATATATGTCGCCCTTCTTAAGAACATGT 6096
Db 145 TTTTATGATGATCTGAGATTTGCTGTGTATATATGTCGCCCTTCTTAAGAACATGT 86
Qy 6097 TGCATTAATGCTCTTGTGATTAATGTTGATTTGACAACTGATTTTAATAACATTTGA 6156
Db 85 TGCATTAATGCTCTTGTGATTAATGTTGATTTGACAACTGATTTTAATAACATTTGC 26
Qy 6157 CT 6158
Db 25 CT 24

RESULT 4
Bg974884 776 bp mRNA linear EST 12-JUN-2001
LOCUS 602842771P1 NCI_CGAP_Mam4 Mus musculus cDNA clone IMAGE:4978744 5',
DEFINITION mRNA sequence.
ACCESSION Bg974884
VERSION Bg974884.1 GI:14362521
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS NIH-MGC
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-rc@mail.nih.gov
            Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth
            Ph.D.
            CNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLM10975 row: m column: 17
            High quality sequence stop: 753.
            location/Qualifiers
                1..776
                /organism="Mus musculus"
                /strain="NMRI"
                /db_xref="taxon:10090"
                /clone="IMAGE:4978744"
                /clone_1ib="NCI CGAP Mam4"
                /tissue_type="tumor, gross tissue"
                /dev_stage="5 months"
FEATURES
SOURCE

```

```

/lab host="HD10B"
/Note="Organ: Mammmary; Vector: pCMV-SPOK6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Priscilla Furth,
NIH Reference for transgenic model: Li et al., Cell Growth
and Differentiation 7, 3-11 (1996)."
```

BASE COUNT 278 a 181 c 216 g 101 t

ORIGIN

Query Match 9.3%; Score 571; DB 13; Length 776;
Best Local Similarity 86.9%; Pred. No. 1.3e-118;
Matches 677; Conservative 0; Mismatches 90; Indels 12; Gaps 4;

```

Qy 327 GAAAAAGCCGAGTTGAAGAAATGCGCGCCACTCCACAGCTGTTGAAGATATCAAAAG 386
Db 1 GAAAAAGCCGAGTTGAAGAAATGCGCGCCACTCCACAGCTGTTGAAGATATCAAAAG 60
Qy 387 GATGGCAGAGAGAGACATCGAAATATGACAGATCCCTGCTTCAGAAAACATGTG 446
Db 61 GACGACGAGAGAGAAACCGGAAATATGACAGATCCCTGCTTCAGAGACAAATGTG 120
Qy 447 GAAGAAATGTAACGCTGCTGAGTCCAGGCTAATGATGTTGCTTCAGAAAGTATTT 506
Db 121 GAAGAAATGCGCAGGCTGCTGAGTCCAGCTAATGACGTCGCTTCAGAAAGTATTT 180
Qy 507 AAATTTGTTGTTTAAATTCACGCTGAAGAAAGATATAAATGAAAGTCATATCTGC 566
Db 181 AAATTTGTTGTTTAAATTCACGCTGAAGAAAGATATAAATGAAAGTCATATCTGC 240
Qy 567 CAATCTACTGCTGCAAGAAAGATGAAGGCGAGAGGCGAGAGGCTGTGCGAGCTTGA 626
Db 241 CAGTACTCACTGTCAGAAAGATGAAGGCGAGAGGCGAGAGGCTGTGCGAGCGAGA 300
Qy 627 GACCAACAGAGGCCAGTGTGAGAGCTGCCGTGCGAGAGTCAGATCCAAAGAAATGTAG 686
Db 301 GACCAACAGAGGCCAGTGTGAGAG--CGTGGCGAATCAGCATCCAAAGAAATGTAG 357
Qy 687 CTGAAGCAATCCACAGAGAGAGAGGACCTGGAAGCAAGAAACAGACAGACACAA 746
Db 358 CTGAAGCAATCCACAGAGAGAGAGGACCTGGAAGCAAGACAGACAGACACAA 417
Qy 747 ATCCCCCTTCAAGCCGATCTGATCAAGCGGCTGAGAGAAAGCCAAAGATGAAGAA 806
Db 418 ATCCCCCTTCAAGCCGATCTGATCTGATGAAGACCGAGAAAGACGCCAAAGATGAAG 477
Qy 807 GAAAAACAAGAGAAAGAGCCCAAGTCCCAAGATCCCGAGAGGCCAGTCAACAGT 866
Db 478 GAAAAACAAGAGAAAGAGCCCAAGTCCCAAGATCTCCGACACAGCCCTGTCAAGCAT 537
Qy 867 GAGACAAATCTTCTTCAAGAAATTTCTTCACTCAAGTGGGCTGGCTGGCGCAAGAG 926
Db 538 GAGACAAATCTTCTTCAAGAAATTTCTTCACTCAAGTGGGCTGGCGCGCAAGAG 597
Qy 927 ACCAGCTTCAAGAAATCAAAAAGAGATGATCGAAACTGCGGAGAGAAAGAGAGCA 986
Db 598 ACCAGCTTCAAGAAATCAAAAAGAGATGATCTGAAACTTCCGAGAAAGACAGAGAGCA 657
Qy 987 GAGCGAGAAAAATGAGACGAGAGAAAAAG---AAAAAGACAGAGCCAGCTCGAGAG 1043
Db 658 GAGCGTGAAGAAATGAGACGAGAGAAAAAGGAGCAAGAGACAGAGCCAGCCGAG 717
Qy 1044 CAGGAG-----CCGCGAGAGACACAGAGCCAGGCTTG-TAGCAGATCTACAGAA 1096
Db 718 GAGCAGAGAGCCTGCGAGAGAGGCGACAGAGCCAGGCGCAGAGATGCCAGCATATGAGAA 776

RESULT 5
BC015830 1269 bp mRNA linear HTC 17-OCT-2001
LOCUS BC015830
DEFINITION Homo sapiens, clone IMAGE:4691862, mRNA.
ACCESSION BC015830
VERSION BC015830.1 GI:16198369

```

KEYWORDS HTC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
AUTHORS 1 (bases 1 to 1269)
TITLE Strausberg, R.
JOURNAL Direct Submission
SUBMITTED (15-OCT-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: CLONTECH
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
contact: amadansystemsbiology.org
Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia Greene, Mark Kettman and Anuradha Madan
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 38 Row: k Column: 11
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4826749
This clone has the following problem: frame shifted.
FEATURES
source Location/Qualifiers
1. 1269
/organism="Homo sapiens"
/db xref="taxon:9606"
/clone="IMAGE:4691862"
/tissue type="Lung"
/clone_lib="NIH_MGC_77"
/lab_host="DH10B"
/note="Vector: pDNR-LIB"
BASE COUNT 419 a 311 c 365 g 174 t
ORIGIN
Query Match 9.08; Score 552.8; DB 11; Length 1269;
Best Local Similarity 69.8; Pred. No. 1.8e-114;
Matches 894; Conservative 0; Mismatches 287; Indels 99; Gaps 7;
Qy 10 GGAGTAGAAGACCACTAGCCATGGCGCAGGAGTTCCACCGAGCAGCGGAGCCCG 69
Db 2 GGGGGAGAGTAGCGGAGGAGCCATGGCGCGGGAGCTCCACCGAGCGCAGCCCGG 61
Qy 70 AGCAGCCG---GCGGGAGGACACGCGAGCGAGCTGTGCTCAGTGGCCATGGGCCG 126
Db 62 AGCAGCCGCGGAGGAGCTCCACCGCGGCTGAGCCGAGCCCGAGCGGCGGCCCT 121
Qy 127 CAGCTGAAGCTTCGGGA-----GCAGCTGGAGACCCCGCGAGCCCGCGCA 177
Db 122 CGGCCGAGGCGGCGCGAGACACCGCGGACCCCGCATCGCTCGGACCCCGCA 181
Qy 178 CCAAGCTCCACAGAGAAGATGCGCAGCTGTCTGTCAACCGCGGTAGCTGAACAAGAG 237
Db 182 CCAAGCTCCTACAGAGAGATGCTGAGCTGTCCACCATCAATGGCGGTAGCTGAGCAAGATG 241
Qy 238 ATGTCCATGTCCAAAGAGAAACACGAGGG----- 268
Db 242 AGCTCAGCTCCAGGAGGGTGACCTTAATGGCCAGAAAGAGCCCTGAACGGTCAAGGAG 301
Qy 269 -----GCAGGAGGAAGATGCTGTGATGAGGATGTTGGACAGCAGAGTCAG 315
Db 302 CCCTAAACAGCCAGGAGGAAGAAGATCAATTGTACAGAGAGTTGGACAGAGACTCTG 361
Qy 316 AAGATGTACAGAGAAAGACCCAGTTGAGAGAAATGGCGGCCCACTCCACAGCTTTGAG 375
Db 362 AAGATGTGAGCAAGAGACTCCGATAAAGAGATGGCTACTTAAGTCAGCGGTTGTTACG 421

Qy 376 ATATCAAAAGGATGGCAGGAGGAGACATCAGAAATAATTGAACAGATCCCTGCTTCAG 435
Db 422 ACATCAGAGATGATGGCAGGAGGAGACACCCGAAATAATCGAACAGATTCCTTCTTCAG 481
Qy 436 AAACAATGTGGAAGAAATGGTACAGCTCTGAGTCCAGGCTAAATGATGTTGGCTTCA 495
Db 482 AAAGCAATTTAGAGAGAGCTAACCAACCACTGAGTCCCAGGCTAAATGATTTGATTTA 541
Qy 496 AGAAAGTATTAAATTTGTTGGTTTTTAAATTCACGGTGAAGAGGATAAAAAATGAAAGT 555
Db 542 AGAAGGTGTTTAAAGTTTGGCTTTAAATTTCACTGTGAAAAAGGATAAGACAGAGAAGC 601
Qy 556 CAGATCTGTCCAACTACTCTCTGTCGAAGAGATGAAGCGGAAGGGGCGAGACCTCTGT 615
Db 602 CTGACACTGTCCAGCTACTCTGTTGAAGAAAGATGAAGGGGAGGAGCAGCA----- 654
Qy 616 TCGAGCTGTGAGACCCACAGGAGCCAGTGTGGAGACTGCCGTGGAGAGTGCACATCCA 675
Db 655 --GGGGCTGGGACCAACAGGACCCCGAGCTTTGGGCT-----GGAGAGCAGCATCCA 706
Qy 676 AAGAAAGTGTGAGCTGAAGCAATTCACAGAGAGCAAGGACCCCTGAAGCAAGACAGA 735
Db 707 AAGAAACCGAACCCAAACAATCTACAGAGAAACCCGAGAGACCTCGAGCGTGAGCAA 766
Qy 736 GCAGCAGAAATCCCTTCAAGCCGAATCTGATCAAGCGGCTGAGGAAGAGCCAAAG 795
Db 767 GCCACGAGAAATTTCTCCCGCCAGCCGAATCTGGCCAA--GCAGTGGAGGAATGCAAG 823
Qy 796 ATGAAGGAGAGAAACACAGAGAGAGCCACCAAGTCCCGAGATCCCGAGCAGCC 855
Db 824 AGGAAGGAGAGAGAAACAAAGAAAGAACTAGCAAGTCTGCAGAAATCTCCGACTAGTC 883
Qy 856 CAGTCAACAGTGAAGCAACATCTTCTTCAAGAAAGTTCTTCACTCAGCTTGGGCCCGCT 915
Db 884 CCGTACAGTGAACAGGATCAACCTTCAAAAAATTTCTTCACTCAGAGTTGGGCCCGCT 943
Qy 916 GGCGCAAGAGACAGCTTCAAGAAATCAAAAGAGGATGATCTGGAACCTGCCGAGAGA 975
Db 944 GGCGCAAAAAAGACCACTTTCAGGAAGCCGAGGAGGATGAAGTGGAAAGTTCAGAGA 1003
Qy 976 GAAGGAGAGAGGAGGAGAAAGTAGACGAGGAGAAAGAAAGAGAGAGCAGCCCT 1035
Db 1004 AAAGGAGAACAGAGCCAGAAAGATAGACAGAGAAAGAGCGGAAAGGAGAGTTGCT 1063
Qy 1036 CGGAGGA-----GCAGGAGCCGCGAGACAGACAGACAGC 1068
Db 1064 CCGAGAACTGACCGCTCCGAGCAAGCCACCCACAGAGAGCGCGAGAAAGTCCCAAG 1123
Qy 1069 AGCCAGGTTGTGAGCAGACTACAGAGAAAGTGGAGTGCCTTTGGAAAGACAGAGTTGGT 1128
Db 1124 AGCCCGGTTATCAGCTGAATATGAGAAAGTTGAGCTGCCCTCAGAGAGGAGCAAGTCAGTG 1183
Qy 1129 ACCTGGAGGATCGTCAGAGAGAGAGTGTCTCTTGGCAACGGAAGTGTTCATGAGA 1188
Db 1184 GCTCGAGGAGACCTTCTGAAGAGAGAAACCTGCTCGTTGGCGACAGAGTGTTCATGAAA 1243
Qy 1189 AGATGGAAGCCCAAGAA 1208
Db 1244 AAAAAAAAAAAAAAAAAA 1263

RESULT 6
BM230239/C
LOCUS BM230239 746 bp mRNA linear EST 31-JAN-2002
DEFINITION K0292D03-3 NIA Mouse Unfertilized Egg cDNA Library (Long) Mus musculus cDNA clone K0292D03 3', mRNA sequence.
ACCESSION BM230239
VERSION BM230239.1 GI:17793419
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

QY	5650	CAG---ATAACATTCGCGGAAAGAAAGACAACTCTTTTAAAGTTTACTGATGCTTA	5705
Db	506	CAGAGAAATACATTCCTCGCAGAGAAATACAGTCTC-TTTTTAAAGTTTACTGATGCTTA	448
QY	5706	GATCGTGGGCGCTTCAGTCCCTGAGAAAGTGTTGTTTTCTTATGACAGCGAGCTCAGAA	5765
Db	447	ACTGTGAGGCGCTGAGTCCCTCTGAAAGTGATGTTTTTCTTATGACAGTGAAGCTCAGAA	388
QY	5766	ATTAAGAACCCCATTTTGGAAATCCAGAGATGTCCTCAATATTTACATGATTTTTTCCCCC	5825
Db	387	ATTAAGAACCCCATTTTGGAGATTCAGAAATGTCCTCAATATTTACATGATTTTTTCCCCC	329
QY	5826	TTTTTTCATATCCAGTCCAGGTTGAAAGAGTCCCTCTGTCAGATTTAAAGCCCTGTC	5885
Db	328	TTTTTTCATATCCAGTCCAGGTTGAAAGAGTCTCTTCTTATGTCAGATTTAAAGCCCTGTC	269
QY	5886	TCTTAATGATATGAGCAAAATGATGTGCTTAAGGCCATGAGATGTTTCTTAATGACAGAG	5945
Db	268	TCTTAATGATATGAGCAAAATGATGTGCTC-ATGCGCATGAGCTATTTCTTAATGACAGAG	210
QY	5946	GAAATCTGTTGACGTTTTTTTGATTTGATCTCTTCTTAATGCTGAGCCGAATCATATGAGA	6005
Db	209	GAAATCTGTTGAT-TACTTTTTGATTTGATCTCTTCTTAATGCTGAGCCGAATCATATGAGA	151
QY	6006	TGCAAGGAGAGCCGCTTCTTACAGATGATTTTGTATGATGATCTGAGTGTGCTGCT	6065
Db	150	TTGAAGTGAAGACCTGCTCTTTACAGATGATTTTGTATGATGATCTGAGTGTGCTGAT	91
QY	6066	TATATCTGTGCCCCCTTC-TTTAAGAACAAATGTCATTTATTTCTTTGATTAATTTGTG	6124
Db	90	TATATCTGTGCCCCCTTC-TTTAAGAACAAATGTCATTTATTTCTTTGATTAATTTGTG	31
QY	6125	ATTTGACAACTGATTTAAATAACATATTT	6154
Db	30	ATTTGACAACTGATTTCAATAAATAATTT	1
RESULT 7			
LOCUS	A1233818/c	563 bp	mRNA linear EST 31-JAN-1999
DEFINITION	EST230506 Normalized rat lung, Bento Soares Rattus sp. cDNA clone		
ACCESSION	A1233818		
VERSION	A1233818.1	GI:3817698	
KEYWORDS	EST.		
SOURCE	Rattus sp.		
ORGANISM	Rattus sp.		
REFERENCE	1 (bases 1 to 563)		
AUTHORS	Lee,N.H., Glodet,A., Chandra,I., Mason,T.M., Quackenbush,J.,		
TITLE	KerLavage,A.R. and Adams,M.D.		
JOURNAL	Rat Genome Project: Generation of a Rat EST (RESE) Catalog & Rat		
COMMENT	Gene Index Unpublished (1998) Other ESTs: TC58449 Contact: Lee, NH The Institute for Genomic Research 9712, Medical Center Drive, Rockville, MD 20850, USA Tel: (301)-838-3529 Fax: (301)-838-0208 Email: nhlee@ligr.org Seq primer: M13-21. Location/Qualifiers 1..563 /organism="Rattus sp." /db_xref="ATCC (inhost):2039638" /db_xref="taxon:10118" /clone="RLUCR55" /clone_lib="Normalized rat lung, Bento Soares" /note="Organ: lung; Vector: pT7T3pac; Site_1: EcoRI; Site_2: NotI"		

BASE COUNT	191 a	107 c	110 g	155 t	
ORIGIN					
Query Match	8.9%; Score 550; DB 9; Length 563;				
Best Local Similarity	99.1%; Pred. No. 7.2e-114;				
Matches	553; Conservative 0; Mismatches 5; Indels 0; Gaps 0;				
QY	5601	ATAGCGAGCTTTGAAAGCTCGAAGCTCAAGCTCAACTGTGTAACTGTGAGCGAAGCAAGATAACATT	5660		
Db	563	AAAGCGAGCTTTGAAAGCTCGAAGCTCAAGCTCAACTGTGTAACTGTGAGCGAAGCAAGATAACATT	504		
QY	5661	CTGCGAAGAGAGACAGAGCTCTTTTAAAGTTTAACTGTGTAGCTGTAGATCTGTGGGCTTCT	5720		
Db	503	CTGCGAAGAGAGAGACAGAGCTCTTTTAAAGTTTAACTGTGTAGCTGTAGATCTGTGGGCTTCT	444		
QY	5721	AGTCCTCTGAAAGTGTGTCTTCTATGACAGCGAGCTCAGAAATAAAGCAAGCTT	5780		
Db	443	AGTCCTCTGAAAGTGTGTCTTCTATGACAGCGAGCTCAGAAATAAAGCAAGCTT	384		
QY	5781	TGAACATCCAGGATGTCCTCAATATACCATGATTTTTCCTCCCTTTTTCCTAATCCAG	5840		
Db	383	TGAACATCCAGGATGTCCTCAATATACCATGATTTTTCCTCCCTTTTTCCTAATCCAG	324		
QY	5841	TCCAGTTTGGAAAGAGTCTCTCTGTGTAGATTAAGCTGTCTCTTAATGATATGGA	5900		
Db	323	TCCAGTTTGGAAAGAGTCTCTCTGTGTAGATTAAGCTGTCTCTTAATGATATGGA	264		
QY	5901	CAATAGATGTGCTCAAGGCCATGAGATGTTTCTTAATGCAAGAGGAATCTGTTGTACGT	5960		
Db	263	CAATAGATGTGCTCAAGGCCATGAGATGTTTCTTAATGCAAGAGGAATCTGTTGTACGT	204		
QY	5961	TTTTTTCATGTACTCTCTATGCTGACCGAATTCATATGACAGATCGAAGTGTCTG	6020		
Db	203	TTTTTTCATGTACTCTCTATGCTGACCGAATTCATATGACAGATCGAAGTGTCTG	144		
QY	6021	TTCTTTACAGATGATTTTTCATGATGACTGGAGTGTCTGTGTATATCTGTGCCCT	6080		
Db	143	TTCTTTACAGATGATTTTTCATGATGACTGGAGTGTCTGTGTATATCTGTGCCCT	84		
QY	6081	TCCTTAAAGAACATGTTGTCATTATGTTCTTTCCTTGGATAAATGTCATTTGACAACTGATTT	6140		
Db	83	TCCTTAAAGAACATGTTGTCATTATGTTCTTTCCTTGGATAAATGTCATTTGACAACTGATTT	24		
QY	6141	AAATAAACATATTGACT	6158		
Db	23	AAATAAACATATTGCTCT	6		
RESULT 8					
BI652823					
LOCUS	603298753F1 NCI_CGAP_Mam3 Mus musculus cdna clone IMAGE:5339244 5',				
DEFINITION	mRNA sequence.				
ACCESSION	BI652823				
VERSION	BI652823.1				
KEYWORDS	EST.				
SOURCE	house mouse.				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
TITLE	1 (bases 1 to 661)				
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/.				
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)				
	Unpublished (1999)				
	Contact: Robert Strausberg, Ph.D.				
	Email: cgapbs-@email.nih.gov				
	Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.				
	cDNA Library Preparation: Life Technologies, Inc.				
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)				
	DNA Sequencing by: Incyte Genomics, Inc.				
	Clone distribution: MGC clone distribution information can be				
	found through the I.M.A.G.E. Consortium/LLNL at:				
	http://image.llnl.gov				
Plate: LLAM1861	row: j column: 13				
High quality sequence stop: 661.					
Location/Qualifiers					
1. .661					
/organism="Mus musculus"					
/strain="129, C57BL/6J, FVB/N"					
/db_xref="taxon:10090"					
/clone="IMAGE:5339244"					
/clone_lib="NCI CGAP Mam3"					
/tissue_type="tumor, gross tissue"					
/dev_stage="10 months"					
/lab_host="DH10B"					
/note="Organ: mammary; Vector: pCMV-SPORT6; Site: 1: SalI;					
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.					
Library constructed by Life Technologies. Investigators					
providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH					
Reference for transgenic model: Xu et al., Nature Genetics					
22, 37-43 (1999)." 98 t					
BASE COUNT	183 a	179 c	201 g		
ORIGIN					
Query Match	8.9%; Score 547.6; DB 13; Length 661;				
Best Local Similarity	90.2%; Pred. No. 2.6e-113;				
Matches	597; Conservative 0; Mismatches 64; Indels 1; Gaps 1;				
QY	2656	CAGTCATTGATGGACAGGGCAGTCCACAGTGTGCGAAGAGCGGTCTCTCTCGTGGATAT	2715		
Db	1	CGGTCTATTGATGGACCA-GGCAGTCACCAAGTCCGAAGAGCGGTCTCTCTCGTGGATAT	59		
QY	2716	CCGCTTCCTGTAACAGACCTCTTGAACACACAGGGGAGAGCCATCCACCTGTTGAAG	2775		
Db	60	CTGCTTCCATGACAGAACCTCTTGAACACACAGGGGAGTGGCCACACCCCTGTTGGAG	119		
QY	2776	AGGTCTACTGAAAAAGACATCATTTGCAGAGAAACTCTCTGTGCTCACCCAGACGTTACGAG	2835		
Db	120	AGGTCTACTGAAAAAGACATCATTTGCAGAGAAACTCTCTGTGCTCACCCAGACGTTACGAG	179		
QY	2836	AGGTCTACTGAAAAAGACATCATTTGCAGAGAAACTCTCTGTGCTCACCCAGACGTTACGAG	2895		
Db	180	GGGGCAAGATGCCCATGACGACATAGTCACCAAGTGAAGTGGATTTTACCTCAGAACGAG	239		
QY	2896	TGACAGCCACAGACCTCAGAGGCTCTCGTACTGAGAAAGTTCACGAGACATCGGGGG	2955		
Db	240	TGACAGCCGAGAAACCAACAGAGGCGCTCCGCGCTGAAGAACTTACCGAAGCATCAGGGG	299		
QY	2956	CCGAAGAGACACAGACATGCTGTCCGACGTTTCCAGCTGACTCTCCAGACACCA	3015		
Db	300	CAGAAGACACAGACATGCTGTCCGACGTTTCCAGCTGACTCTCCAGACACCA	359		
QY	3016	CAGAGGAAGCCACCCAGTTTCCAGAGGTAGAGATGCTGTGCTAGATACAGAAAGAGG	3075		
Db	360	CAGAGGAAGCCACCCAGTTTCCAGAGGTAGAGATGCTGTGCTAGATACAGAAAGAGG	419		
QY	3076	AGGCGCAGACGAGGCGATCTCTCAAGCCCTTCAGACAAAGGTGAAGAGAGTCCAGG	3135		
Db	420	AGGCGCAGACGAGGCGCTCTCCAAGCCCTTCAGACAAAGGTGAAGAGAGTCCAGG	479		
QY	3136	TGCTGCAACCCAGACATGTGTCAGAGAACCGGGTCAAAAGCACTGGAGAGGTTGAGGAGG	3195		
Db	480	TGCTGCAACCCAGACATGTGTCAGAGAACCGGGTCAAAAGCACTGGAGAGGTTGAGGAGG	539		
QY	3196	TAGAGGAGGACTCCGAGTCTCTGGCTTTCGAGAAAGAGAGGAGCTTTATCCGAAAGGAC	3255		
Db	540	TAGAGGAGGACTCCGAGTCTCTGGCTTTCGAGAAAGAGAGGAGCTTTATCCGAAAGGAC	599		
QY	3256	CCGTGCGAAGCTGAGGCTGAGCATCTTCCACAGGCTCTGAGACTGAGAGGCTACTC	3315		
Db	600	CCGTGCGAAGCTGAGGCTGAGCATCTTCCACAGGCTCTGAGAGGCTCTGAGAGGCTACTC	659		
QY	3316	CA	3317		
Db	660	CA	661		

RESULT 9
 BM119923/c 710 bp mRNA linear EST 01-FEB-2002
 L0932B09-3 NIA Mouse Newborn Kidney CDNA Library (Long) Mus
 DEFINITION musculus cDNA clone L0932B09 3', mRNA sequence.
 ACCESSION BM119923
 VERSION BM119923.1 GI:17087949
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 710)
 Pao, Y., Kargul, G.J., Dudekula, D.B., Qian, Y., Pantano, S., Lim, M.K.
 and Ko, M.S.H.
 TITLE Systematic Analyses of NIA Mouse Newborn Kidney CDNA Library
 JOURNAL Unpublished (2001)
 COMMENT Contact: Dawood B. Dudekula
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Cassel Drive, Suite 4000, Baltimore, MD 21224-6820, USA
 Email: cdna@igsun.grc.nia.nih.gov
 Plate: L0932 row: E column: 09
 Seq primer: -21M13 Forward
 High quality sequence stop: 710
 POLYA=Yes.

FEATURES
 source
 Location/Qualifiers
 1..710
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="tastrain:L0932B09-3"
 /db_xref="taxon:10090"
 /clone="L0932B09"
 /clone_lib="NIA Mouse Newborn Kidney cDNA library (Long)"
 /tissue_type="Newborn Kidney"
 /dev_stage="Newborn"
 /lab_host="DH10B"
 /note="Organ: kidney; Vector: pSPORT1 (Invitrogen); Site 1: SalI; Site 2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (http://igsun.grc.nia.nih.gov/cDNA). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11:1553-1558 (2001). [PMID:1154199]). In brief , double-stranded cDNAs were synthesized with an Oligo(dT) primer (Invitrogen: 5'-GACTAGTCTCAGATCGGAGCGGCCGCTTTT-3') from 26 microgram of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to lone-linker Lf-SalI, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer SalI-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes, and cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with ligation mixture by the standard chemical method. The average insert size is about 3.0 kb. The library was constructed by Yulan Pao (NIA)."

BASE COUNT 252 a 130 c 137 g 191 t
 ORIGIN

Query Match 8 6%; Score 532.2; DB 13; Length 710;
 Best Local Similarity 90.5%; Pred. No. 8e-110;
 Matches 636; Conservative 0; Mismatches 58; Indels 9; Gaps 6;

OY 5459 CCACCTTCAACTGAGAGATCCTCTTATGATATATATATGTTATGAGTCT 5518
 |||||
 DB 710 CCACCTTCAACTGAGAGATCCTCTTATGATATATATGATGTTATGAGTCT 651
 |||||

OY 5519 CCTCTGTACTATTGTATA-TTTTCTCTAAAGTTAAGACATGCTTTTGTATATG 5577

DB 650 CCTCTGTACTATTGTATA-TTTTCTCTAAAGTTAAGACATGCTTTTGTATATG 591
 |||||

OY 5578 CAATATATAACGGGTGTGACGCCATAGCAGCGCTTTGAAAAAGCTCAAGCCTCACTGTA 5637
 |||||

DB 590 CAGATATATAATGGGTGTGACGCCATAGCAGCGCTTTGAAAAAGCTCAAGCCTCACTGTA 531
 |||||

OY 5638 ACCCTGACCAACG---ATAACATTCCTGGCAAGAAAGACAAAGCTTTTAAAGTT 5693
 |||||

DB 530 ACCGTGACCAACGAAAGAAATTAACATTCCTGGCAAGAAAGCTTTTAAAGTT 472
 |||||

OY 5694 TACGATGCTTAGATCTGTGGGCTTCTAGTCCCTGAAAGGTGTGTTTCCATGACGA 5753
 |||||

DB 471 TACGATGCTTAGATCTGTGGGCTTCTAGTCCCTGAAAGGTGTGTTTCCATGACGA 412
 |||||

OY 5754 GCGAGCTCAGAAATTAACCCCATTTTGAACATCCAGAGTGTCCCATATATACATGA 5813
 |||||

DB 411 GTGAGCTCAGAAATTAACCCCATTTTGAACATCCAGAGTGTCCCATATATACATGA 352
 |||||

OY 5814 TTTTTCCTCCCTTTTTCCTATATCCAGTCCAGTTGAAAGAAATCTCTCTGTGACGA 5873
 |||||

DB 351 CATTTTCTTTTCTTTCTTAAATCAGTCCAGTTGAAAGAAATCTCTCTGTGACGA 292
 |||||

OY 5874 TTAAGCCCTGTCTCTATATGATATGACAAATGAGTGTGCTAAGGCCATGAGATGTTTC 5933
 |||||

DB 291 TTAAGCCCTGTCTCTATATGATATGACAAATGAGTGTGCTAAGGCCATGAGATGTTTC 233
 |||||

OY 5934 CTAATGACAAAGAAATCTGTGACGTTTGTGATGTAATCTCTCTATGCTGACCGAA 5993
 |||||

DB 232 CTAATGACAAAGAAATCTGTGACGTTTGTGATGTAATCTCTCTATGCTGACCGAA 174
 |||||

OY 5994 TTCAATGACAGATGAAAGTGAAGTCTGTTCTTTACAGATGATTTTGATGATGATGCA 6053
 |||||

DB 173 TTCAATGACAGATGAAAGTGAAGTGAAGTCTGTTCTTTACAGATGATTTTGATGATGATGCA 114
 |||||

OY 6054 GTTGTGCTGTATATATCTGTGCCCCCTTC--TTTAAAGCAATGTGATATGTTCCCTT 6112
 |||||

DB 113 GTTGTGCTGTATATATCTGTGCCCCCTTTTAAAGCAATGTGATATGTTCCCTT 54
 |||||

OY 6113 GGATTAATTTGATTTGACAACTGATTTAAATTAACATTTTGG 6155
 |||||

DB 53 GGATTAATTTGATTTGACAACTGATTTCAATATAAATATTTGG 11
 |||||

RESULT 10
 BS546856 668 bp mRNA linear EST 26-OCT-2001
 LOCUS BS546856
 DEFINITION musculus cDNA full-length enriched, 0 day neonate eyeball Mus
 mRNA for SSeCKS, mRNA sequence.
 ACCESSION BS546856
 VERSION BS546856.2 GI:16447501
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 668)
 Arakawa, T., Caminini, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.,
 Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J., Kono, H., Kouda,
 M., Koye, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M., Okazaki,
 Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki,
 D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H.,
 Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T.,
 Muramatsu, M. and Hayashizaki, Y.
 RIKEN Mouse ESTs (Arakawa, T., et al. 2001)
 TITLE Unpublished (2001)
 JOURNAL On Jul 31, 2000 this sequence version replaced gi:9618284.
 COMMENT Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

[illegible]

Db 249 ATGAGTGGCC-ATGCGCATGAGCTATTTCCTAAGCAGAAAGATTTGTTGT-TACTTT 192
 QY 5964 TTGATTTACTCTTTCTATGCTGAGCCGAATTCATATGCAATGCAAGTGAAGTCTGTTTC 6023
 Db 191 TTGGATTCTACTCTCTCTAGAGCTGAGCCGAATTCATATGCAATGCAAGTGAAGTCTGTTTC 132
 QY 6024 TTTCAGATGATTTTGTATGATATCTGAGAGTTTCTGATTTATATCTGTCGCCCTTC- 6082
 Db 131 TTTCAGATGATTTTGTATGATATCTGAGAGTTTCTGATTTATATCTGTCGCCCTTC 72
 QY 6083 TTTCAGATGATTTTGTATGATATCTGAGAGTTTCTGATTTATATCTGTCGCCCTTC 6142
 Db 71 TTTCAGATGATTTTGTATGATATCTGAGAGTTTCTGATTTATATCTGTCGCCCTTC 12
 QY 6143 ATTAACATATT 6153
 Db 11 ATTAACATATT 1

RESULT 14
 A1103489 476 bp mRNA linear EST 08-JAN-1999
 LOCUS EST212778 Normalized rat embryo, Bento Soares Rattus sp. cDNA clone
 DEFINITION REMCC63.3 end, mRNA sequence.
 ACCESSION A1103489
 VERSION A1103489.1 GI:4134176
 KEYWORDS EST.
 SOURCE Rattus sp.
 ORGANISM Rattus sp.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 476)
 Lee, N.H., Glodek, A., Chandra, I., Mason, T.M., Quackenbush, J.,
 Kerkavag, A.R. and Adams, M.D.
 Rat Genome Project: Generation of a Rat EST (RST) Catalog & Rat
 Gene Index
 Unpublished (1998)
 On Oct 6, 1998 this sequence version replaced gi:3708028.
 Contact: Lee, NH
 The Institute for Genomic Research
 9712, Medical Center Drive, Rockville, MD 20850, USA
 Tel: (301)-838-3529
 Fax: (301)-838-0208
 Email: nhlee@igr.org
 Seq primer: M13-21.

FEATURES
 source
 1..476
 /organism="Rattus sp."
 /db_xref="taxon:10118"
 /clone="REMCC63"
 /clone_1ib="Normalized rat embryo, Bento Soares"
 /dev_stage="embryo 8, 12, 18 dpc"
 /note="Vector: pT713pac, Site_1: EcoRI, Site_2: NotI"
 BASE COUNT 170 a 94 c 90 g 122 t
 ORIGIN

Query Match 7.6%; Score 467.8; DB 9; Length 476;
 Best Local Similarity 9.6%; Pred No. 3.1e-95;
 Matches 469; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 5688 AAGTTTACTGATGCTAGTCTGAGGCTTCTAGTCTCTGGAAGTGTGTTTCTTA 5747
 Db 476 AAGTTTACTGATGCTAGTCTGAGGCTTCTAGTCTCTGGAAGTGTGTTTCTTA 417
 QY 5748 TGCACAGCGAGCTCAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5807
 Db 416 TGCACAGCGAGCTCAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 357
 QY 5808 CCAATGATTTTCTCCCTTTTGTCTATCCAGTCCAGTGTGGAAGAAAGTCTCTGT 5867
 Db 356 CCAATGATTTTCTCCCTTTTGTCTATCCAGTCCAGTGTGGAAGAAAGTCTCTGT 297

QY 5868 GTCAGATTAAGCCCTCTCTCTATGATGATGACAATATGAGTGTGCTTAAGCCATGAGA 5927
 Db 296 GTCAGATTAAGCCCTCTCTCTATGATGATGACAATATGAGTGTGCTTAAGCCATGAGA 237
 QY 5928 TGTTCCTAATGACAGAAAGAAATCTGTTGACGTTTGTGATGATCTTCTTAATGCTGG 5987
 Db 236 TGTTCCTAATGACAGAAAGAAATCTGTTGACGTTTGTGATGATCTTCTTAATGCTGG 177
 QY 5988 ACCGAATTCATATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 6047
 Db 176 ACCGAATTCATATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 117
 QY 6048 ACTGAGATTTCTGTTATATCTGAGCCCTCTTAAAGAACATGTCATTATGT 6107
 Db 116 ACTGAGATTTCTGTTATATCTGAGCCCTCTTAAAGAACATGTCATTATGT 57
 QY 6108 CTTTGGATTAATTTGATTTGACACTGATTAAATTAACATATTGACT 6158
 Db 56 CTTTGGATTAATTTGATTTGACACTGATTAAATTAACATATTGACT 6

RESULT 15
 B6670384 573 bp mRNA linear EST 30-APR-2001
 LOCUS DRNBAB11 Rat DRG Library Rattus norvegicus cDNA clone DRNBAB11 5',
 DEFINITION mRNA sequence.
 ACCESSION B6670384
 VERSION B6670384.1 GI:13892483
 KEYWORDS EST.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 573)
 Xiao, H.S., Huang, Q.H., Zhang, F.X., Bao, L., Lu, Y.J., Guo, C., Yang, L.,
 Huang, W.J., Fu, G., Xu, S.H., Cheng, X.P., Yan, Q., Zhu, Z.D., Zhang,
 X., Chen, Z., Han, Z.G. and Zhang, X.
 Identification of gene expression profile of dorsal root ganglion
 in the rat peripheral axotomy model of neuropathic pain
 Proc. Natl. Acad. Sci. U.S.A. 99 (12), 8360-8366 (2002)
 22056133
 Contact: Zhang Xu
 Laboratory of Sensory System
 Institute of Neuroscience
 320 Yue Yang Road, Shanghai 200031, P.R.China
 Tel: 86-21-64748700-121
 Fax: 86-21-64713446
 Email: xu.zhang@ion.ac.cn
 This clone is also available at Chinese National Human Genome
 Center at Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park,
 Pudong New Area, P.R.China. Please contact with Zhang Xu
 (xu.zhang@ion.ac.cn) or Han Zeguang (hanzg@shc.sh.cn)
 PCR primers
 FORWARD: T3
 BACKWARD: T7
 Seq primer: T3
 POLYA=NO.

FEATURES
 source
 1..573
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="DRNBAB11"
 /clone_1ib="Rat DRG Library"
 /sex="male"
 /tissue_type="dorsal root ganglion"
 /dev_stage="adult"
 /note="Total RNA was isolated from hypothalamus and
 transcribed into cDNA, which was then used as template in
 PCR. The PCR products were separated on sequencing gel.
 The differential bands were cut, reamplified, cloned into
 pMD18-T vector and confirmed by Northern blot."

Query Match 7.6%; Score 467.8; DB 9; Length 476;
 Best Local Similarity 9.6%; Pred No. 3.1e-95;
 Matches 469; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 5688 AAGTTTACTGATGCTAGTCTGAGGCTTCTAGTCTCTGGAAGTGTGTTTCTTA 5747
 Db 476 AAGTTTACTGATGCTAGTCTGAGGCTTCTAGTCTCTGGAAGTGTGTTTCTTA 417
 QY 5748 TGCACAGCGAGCTCAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5807
 Db 416 TGCACAGCGAGCTCAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 357
 QY 5808 CCAATGATTTTCTCCCTTTTGTCTATCCAGTCCAGTGTGGAAGAAAGTCTCTGT 5867
 Db 356 CCAATGATTTTCTCCCTTTTGTCTATCCAGTCCAGTGTGGAAGAAAGTCTCTGT 297

BASE COUNT	151 a	131 c	115 g	170 t	6 others
ORIGIN					
Query Match 7.5%; Score 460.2; DB 12; Length 573;					
Best Local Similarity 88.0%; Pred. No. 1.7e-93;					
Matches 557; Conservative 0; Mismatches 14; Indels 62; Gaps 3;					
Qy 5088	GAATCCTAAGATGTAGTGTGCTGATTTGATCTGTGAAGACCAAGTGTGAACAAGTC	5147			
Db 1	GAATCCTAAGATGTAGTGTGCTGATTTGATCTGTGAAGACCAAGTGTGAACAAGTC	60			
Qy 5148	ACAGAACAAAGATGCTGCTGTTGGACCTTGAGACCAAGATTTACAGCCCATGAGATCCA	5207			
Db 61	ACAGAACAAAGATGCTGCTGTTGGACCTTGAGACCAAGATTTACAGCCCATGAGATCCA	120			
Qy 5208	GAGAGCAGGGCCGTCCTCAATGATTTCCACCCAGTAGACACCCCGACAAATTCAGAGGCTTC	5267			
Db 121	GAGAGCAGGGCCGTCCTCAATGATTTCCACCCAGTAGACACCCCGACAAATTCAGAGGCTTC	180			
Qy 5268	ATCGGGAGCTAGACCCAGCTAACATTTCTCGTTTCAAGACTGCCTTTGATTTGCCCTT	5327			
Db 181	ATCGGGAGCTAGACCCAGCTAACATTTCTCGTTTCAAGACTGCCTTTGATTTGCCCTT	198			
Qy 5328	GATGCCGTCCGTGTATTTCTAACAATTTCTCGTTTCAAGACTGCCTTTGATTTGCCCTT	5387			
Db 199	-----CTAACAATTTCTCGTTTCAAGACTGCCTTTGATTTGCCCTT	240			
Qy 5388	GATGCCGTCCGTGTATTTCTCGTTTCAAGACTGCCTTTGATTTGCCCTT	5446			
Db 241	GATGCCGTCCGTGTATTTCTCGTTTCAAGACTGCCTTTGATTTGCCCTT	300			
Qy 5447	CCATACCTAGTTCACACTTCAAACTGGAGCATCCTCTTTATGTATTTATGTATGTT	5506			
Db 301	CCATACCTAGTTCACACTTCTCAAACTGGAGCATCCTCTTTATGTATTTATGTATGTT	360			
Qy 5507	TTATGTAGTCTCTCTGTTACCTATTGTATATTTTCTTAACGTTTAAAGCATGCTT	5566			
Db 361	NTATGTAGNCCCTCTNCTGTACCTATTGTATATTTTCTTAACGTTTAAAGCATGCTT	420			
Qy 5567	TTTGTATTATGCAATATATATACGGGTGTGAGCCATAGCCGCTTTGAAAAGCTCCAAG	5626			
Db 421	TTTGTATTATGCAATATATATACGGGTGTGAGCCATAGCCGCTTTGAAAAGCTCCAAG	480			
Qy 5627	CCTCAACTGTAACTGTCAGCAAAACAGATAACA-TTCCTGGCAAGAGACAGATCTTTT	5685			
Db 481	CCTCAACTGTAACTGTCAGCAAAACAGATAACA-TTCCTGGCAAGAGACAGATCTTTT	540			
Qy 5686	TTAAAGTTTACTGATGCTTAGATCTGTGGCTT	5718			
Db 541	TTAAAGNTACTGATGCTTAGATCTGTGGCTT	573			

Search completed: December 12, 2002, 16:19:41
Job time : 7377 secs

THIS PAGE BLANK (uspro)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2002, 16:39:22 ; Search time 123 Seconds
(without alignments)
1729.008 Million cell updates/sec

Title: US-09-902-432-4
Perfect score: 8073
Sequence: 1 MGRGSTEQRSPQAGSDT.....AWAQRKCLRLQLKAPVSK 1596

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8049	99.7	1596	AAW31347	Rat tumour suppress
2	5387.5	66.7	1346	AAW31346	Rat tumour suppress
3	4334	53.7	1783	ABR97448	Novel human protei
4	4294	53.2	1781	AAQ17365	Human gravin. Hom
5	4274.5	52.9	1780	AAW53863	Human gravin. polyp
6	4274.5	52.9	1780	AAW53863	Human gravin. polyp
7	4262.5	52.8	1795	AAW53862	Human gravin. polyp
8	952.5	11.8	292	AAW53862	Novel human diagno
9	952.5	11.8	292	AAW53862	Novel human diagno
10	656	8.1	651	AAW53875	Gravin polypeptide

11	656	8.1	651	21	ABW18410	Human gravin PKA r
12	642	8.0	2768	22	ABW68397	Drosophila melanog
13	639.5	7.9	453	19	AAW53867	Gravin polypeptide
14	639.5	7.9	453	21	ABW18407	Human gravin PKA r
15	557.5	6.9	6815	22	ABW66811	Drosophila melanog
16	521.5	6.5	396	19	AAW53868	Gravin polypeptide
17	521.5	6.5	396	21	ABW18408	Human gravin PKA r
18	492	6.1	2519	22	ABW16636	Novel human diagno
19	485	6.0	3263	22	ABW67210	Drosophila melanog
20	462	5.7	1786	18	AAW24790	P. falciparum live
21	462	5.7	1787	23	AAU96699	Plasmodium falcipo
22	451	5.6	1468	22	ABW62991	Drosophila melanog
23	440.5	5.5	1812	22	ABW58022	Drosophila melanog
24	431	5.3	3257	22	ABW67502	Drosophila melanog
25	430	5.3	3111	22	ABW60327	Drosophila melanog
26	429	5.3	2665	22	ABW33490	Human peptide #965
27	429	5.3	2665	22	ABW33490	Human peptide #965
28	429	5.3	2665	22	ABW18950	Protein #949 encod
29	429	5.3	2665	22	AAW54270	Human brain expres
30	429	5.3	2665	22	AAW66665	Human bone marrow
31	429	5.3	2665	22	AAW14533	Peptide #967 encod
32	429	5.3	2665	22	AAW26950	Peptide #987 encod
33	429	5.3	2665	22	AAW02259	Peptide #941 encod
34	429	5.3	2665	23	ABG36319	Human peptide enco
35	429	5.3	2748	22	ABW58843	Drosophila melanog
36	429	5.3	3266	21	AAW42491	Human ORFX ORF2255
37	421	5.2	1714	22	ABG02138	Novel human diagno
38	421	5.2	1714	22	ABG06417	Novel human diagno
39	420	5.2	2803	23	ABW08161	Human cytoskeleton
40	414.5	5.1	2622	22	ABG06418	Novel human diagno
41	411	5.1	2618	22	ABG02135	Novel human diagno
42	409.5	5.1	2515	22	ABW12281	Human secreted pro
43	409.5	5.1	2515	22	AAW80268	Human protein SEQ
44	409.5	5.1	2515	22	AAW80268	Human protein SEQ
45	409	5.1	1793	22	ABW59613	Drosophila melanog

ALIGNMENTS

RESULT 1

AAW31347
ID AAW31347 standard; Protein; 1596 AA.

XX AAW31347;

XX AC

XX AAW31347;

XX AC

XX AAW31347;

XX AC

XX AAW31347;

XX AC

XX AAW31347;

XX AC

XX AAW31347;

XX AC

XX AAW31347;

XX AC

XX AAW31347;

XX AC

XX AAW31347;

XX AC

XX AAW31347;

XX AC

XX AAW31347;

XX AC

XX AAW31347;

XX AC

XX AAW31347;

XX AC

XX AAW31347;

XX AC

XX AAW31347;

XX AC

XX AAW31347;

XX AC

XX AAW31347;

XX AC

FT	Modified-site	504..526	site"
FT	/note="proposed protein kinase C phosphorylation		
FT	Modified-site	592..614	site"
FT	/note="proposed protein kinase C phosphorylation		
FT	Modified-site	741..766	site"
FT	/note="proposed protein kinase C phosphorylation		
XX	MO9740059-A1.		
XX	30-OCT-1997.		
XX	18-APR-1997; 97MO-US06830.		
XX	18-JUN-1996; 96US-0665401.		
XX	19-APR-1996; 96US-0635121.		
XX	(GELM/) GELMAN I.		
XX	(JAKE/) JAKEN S.		
XX	Gelman I, Jaken S;		
XX	WPI: 1997-535770/49.		
XX	N-PSDB; AAV02302.		
PT	Tumour suppressor gene SSeCKs - used as a mitotic regulator, and		
PS	inhibitor of malignant phenotype		
PS	Claim 4; Fig 11A-L; 162pp; English.		
XX	This polypeptide comprises the novel tumour suppressor protein		
CC	SSeCKs that is a substrate of protein kinase C and which acts as a		
CC	negative regulatory of mitosis and as an inhibitor of the		
CC	transformed phenotype. Its amino acid sequence was deduced from an		
CC	isolated nucleic acid molecule (see AAV02302). The SSeCKs nucleic		
CC	acid, as well as homologous and hybridising nucleic acids are		
CC	claimed, as are isolated proteins encoded by such nucleic acids,		
CC	vectors comprising the nucleic acids, host cells, and methods of		
CC	inhibiting the expression of a transformed phenotype in a host cell		
CC	by introducing the nucleic acids. Introduction of a SSeCKs nucleic		
CC	acid or gene product into a host cell inhibits mitosis of the host		
CC	cell, allowing the treatment of diseases associated with disorders		
CC	of proliferation and/or with the expression of a malignant		
CC	phenotype. SSeCKs can also be used to treat or identify disorders		
CC	of cytoskeletal structure and cellular architecture (such as		
CC	Alzheimer's disease), and may be a marker for aberrancies in		
CC	fertility and/or nervous system development.		
XX	Sequence 1596 AA:		
XX	Query Match 99.7%; Score 8049; DB 18; Length 1596;		
XX	Best Local Similarity 99.8%; Pred. No. 0;		
XX	Matches 1593; Conservative 0; Mismatches 3; Indels 0; Gaps 0;		
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db	1	1	1
Qy	1	1	1
Db			

Oy	241	LOAESDOAAEEKANDGEKOKKEPTKSPESPSVSNSETTSSPKKFFTHGMAMRKTS	300
Db	241	LOAESDOAAEEKANDGEKOKKEPTKSPESPSVSNSETTSSPKKFFTHGMAMRKTS	300
Oy	301	FKSKEDDLTAEKKEQAEKVDDEBEKKEPTASEEQEPADDTQARLSADYKEVPL	360
Db	301	FKSKEDDLTAEKKEQAEKVDDEBEKKEPTASEEQEPADDTQARLSADYKEVPL	360
Oy	361	EDVOGDLLEASSEKCPPLATEPDEMEHQVNAEYHVSIVTEKTEBEEQGGGAEGGV	420
Db	361	EDVOGDLLEASSEKCPPLATEPDEMEHQVNAEYHVSIVTEKTEBEEQGGGAEGGV	420
Oy	421	VEGGESELPREKLAEOVEQAEAPABELMKSRKCVSGGHOTLTDLSPEKTLPGPE	480
Db	421	VEGGESELPREKLAEOVEQAEAPABELMKSRKCVSGGHOTLTDLSPEKTLPGPE	480
Oy	481	GIVSEVEMLSQERIRIVQSGPLKULFSSSGKLKULSGKKQKRGCGGDEBGEYQH1TE	540
Db	481	GIVSEVEMLSQERIRIVQSGPLKULFSSSGKLKULSGKKQKRGCGGDEBGEYQH1TE	540
Oy	541	SPESADQKESASSPEPEPTTCLKEGPLAPODGAEEGTTSDEBKKEGITTPMASF	600
Db	541	SPESADQKESASSPEPEPTTCLKEGPLAPODGAEEGTTSDEBKKEGITTPMASF	600
Oy	601	KKMTTPKKRRRRESDEKBEELKXKSA1LSTSDTSVEMODEYKTYGEBQKPEEPPRRV	660
Db	601	KKMTTPKKRRRRESDEKBEELKXKSA1LSTSDTSVEMODEYKTYGEBQKPEEPPRRV	660
Oy	661	DTYSWML1CVSSSKKRRKAKASSPDDEGPPTLGDGSHRAEASKDEKATDVAVPASTQ	720
Db	661	DTYSWML1CVSSSKKRRKAKASSPDDEGPPTLGDGSHRAEASKDEKATDVAVPASTQ	720
Oy	721	EODQAGSSSPEPAGSPSBEQGVSTWESFKRLVTPRKKSKSKLEBKADSSVEQLSTEIE	780
Db	721	EODQAGSSSPEPAGSPSBEQGVSTWESFKRLVTPRKKSKSKLEBKADSSVEQLSTEIE	780
Oy	781	PSREESVWS1KKRT1PGRKKRADGQEOQATVDSGPVEINDDPVPVAVPLSEYNAVER	840
Db	781	PSREESVWS1KKRT1PGRKKRADGQEOQATVDSGPVEINDDPVPVAVPLSEYNAVER	840
Oy	841	EKMAQONTTEL1POLLGAVVYSEELSTLTHVTSVAVIDGTRAVTSVEERSPSM1SASYTE	900
Db	841	EKMAQONTTEL1POLLGAVVYSEELSTLTHVTSVAVIDGTRAVTSVEERSPSM1SASYTE	900
Oy	901	PLEHTAGEAMPPEVEYTEKDI1AEETPVL1OTL1PEGKADHDMVTSEVDTSEA1VTATET	960
Db	901	PLEHTAGEAMPPEVEYTEKDI1AEETPVL1OTL1PEGKADHDMVTSEVDTSEA1VTATET	960
Oy	961	SEALRTEEVTBASGABETTDMANSQ1LTDSDTTBEAT1PQOEVESGVLDTBEEROTQA	1020
Db	961	SEALRTEEVTBASGABETTDMANSQ1LTDSDTTBEAT1PQOEVESGVLDTBEEROTQA	1020
Oy	1021	ILQAVAKVKEESQVPA1TOVORTSGALEKXEVEBDESVLASKEKDMWPKPVOEAG	1080
Db	1021	ILQAVAKVKEESQVPA1TOVORTSGALEKXEVEBDESVLASKEKDMWPKPVOEAG	1080
Oy	1081	AENHLAGSE1TGOATPES1LEPEVTADVDHVA1COY1K1QOLMEQAVAPESSET1LTDSETN	1140
Db	1081	AENHLAGSE1TGOATPES1LEPEVTADVDHVA1COY1K1QOLMEQAVAPESSET1LTDSETN	1140
Oy	1141	GSTPLADSD1ADGTQODETT1DSQDSKATA1VAFQSGVTEBEA1TAQKEPST1L1PNNVPAQE	1200
Db	1141	GSTPLADSD1ADGTQODETT1DSQDSKATA1VAFQSGVTEBEA1TAQKEPST1L1PNNVPAQE	1200
Oy	1201	EHGEPGRDVL1EP1QOELTAA1PV1AKTEVQGBGVMLDSEKXKEQEV1FVHSGPNSQ	1260
Db	1201	EHGEPGRDVL1EP1QOELTAA1PV1AKTEVQGBGVMLDSEKXKEQEV1FVHSGPNSQ	1260
Oy	1261	KAAVY1TADSEW1GAVGOEKESTEVOS1LEJEGEMETD1VEKRE1TKPEOV1SEEGEOETA	1320
Db	1261	KAAVY1TADSEW1GAVGOEKESTEVOS1LEJEGEMETD1VEKRE1TKPEOV1SEEGEOETA	1320
Oy	1321	APHEG1YGRV1L1LDMPSSEKALG1SGSPSL1PDODKAG1CE1VQYQSD1T1VTQ1AE	1380

```

Db 1321 APEHEGTYGKPVLTLDMESSRGKALGSLGGSPSLPDDQKAGCIEVQVQSDDTTVTQAE 1380
Qy 1381 AVEKVIETWISGETSEPECYGAHLLPAEKSSATGGHWTLOHAEDTVPLGPESQAESIPI 1440
Db 1381 AVEKVIETWISGETSEPECYGAHLLPAEKSSATGGHWTLOHAEDTVPLGPESQAESIPI 1440
Qy 1441 IVTPAPESTLHPDLOGEISASQSRSEBEDXPADGPDADGKESIAIEKVLKAEPEILELE 1500
Db 1441 IVTPAPESTLHPDLOGEISASQSRSEBEDXPADGPDADGKESIAIEKVLKAEPEILELE 1500
Qy 1501 SKSNKIVLNLQTAVDQPARTETAPETHAYDSQTPACRLDSREPNRCWTMKDKAKMH 1560
Db 1501 SKSNKIVLNLQTAVDQPARTETAPETHAYDSQTPACRLDSREPNRCWTMKDKAKMH 1560
Qy 1561 PVPQPREDLQVLTLEAWAQPCKLPLRLQKAPVSK 1596
Db 1561 PVPQPREDLQVLTLEAWAQPCKLPLRLQKAPVSK 1596

RESULT 2
AAW31346
ID AAW31346 standard; Protein; 1346 AA.
XX AC AAW31346;
XX
XX 20-JUL-1998 (first entry)
XX
XX Rat tumour suppressor protein SSeCKs (active truncated form).
XX
XX SSeCKs; tumour suppressor gene; rat; protein kinase C; mitosis;
XX cancer; malignancy; cell proliferation; Alzheimer's disease;
XX therapy.
XX
XX Rattus sp.
XX
XX Key Location/Qualifiers
FH Misc-difference 72
FT /note= "encoded by TCT"
FT Misc-difference 496
FT /note= "encoded by GTC"
FT Misc-difference 499
FT /note= "encoded by GTC"
FT Misc-difference 757
FT /note= "encoded by CCC"
FT Misc-difference 785
FT /note= "encoded by AGG"
FT Region 24..32
FT /note= "glycine-rich region"
FT Peptide 131..134
FT Region 138..141
FT /note= "nuclear localisation signal"
FT Peptide 202..205
FT /note= "glycine-rich region"
FT Peptide 289..292
FT /note= "nuclear localisation signal"
FT Region 306..315
FT /note= "glycine-rich region"
FT Peptide 368..371
FT /note= "nuclear localisation signal"
FT Region 1013..1035
FT /note= "zinc finger"
XX
XX WO9740059-A1.
XX
XX 30-OCT-1997.
XX
XX 18-APR-1997; 97WO-US06830.
XX
XX 18-JUN-1996; 96US-0665401.
XX
XX 19-APR-1996; 96US-0635121.
XX

```

```

PA (GELM/) GELMAN I.
PA (JAKE/) JAKEN S.
XX
XX Gelman I, Jaken S;
XX
XX WPI; 1997-535770/49.
XX N-PSDB; AAV02301.
XX
XX Tumour suppressor gene SSeCKs - used as a mitotic regulator, and
XX inhibitor of malignant phenotype
XX
XX Example 6; Fig 3A-G; 162pp; English.
XX
XX This polypeptide comprises an active truncated form of the novel
XX tumour suppressor protein SSeCKs that is a substrate of protein
XX kinase C and which acts as a negative regulatory of mitosis and as
XX an inhibitor of the transformed phenotype. Its amino acid sequence
XX was deduced from a partial cDNA clone (see AAV02301) obtained from
XX NIH3T3 cells. The full-length SSeCKs sequence (see AAW31347) is also
XX provided. Full-length SSeCKs nucleic acid (see AAV02301), as well as
XX homologous and hybridising nucleic acids are claimed, as are
XX the nucleic acids, host cells, and methods of inhibiting the
XX expression of a transformed phenotype in a host cell by introducing
XX the nucleic acids. Introduction of a SSeCKs nucleic acid or gene
XX product into a host cell inhibits mitosis of the host cell,
XX allowing the treatment of diseases associated with disorders of
XX proliferation and/or with the expression of a malignant phenotype.
XX SSeCKs can also be used to treat or identify disorders of
XX cytoskeletal structure and cellular architecture (such as
XX Alzheimer's disease), and may be a marker for aberrancies in
XX fertility and/or nervous system development.
XX
XX Sequence 1346 AA;
XX
XX Query Match 66.7%; Score 5387.5; DB 18; Length 1346;
XX Best Local Similarity 91.1%; Pred. No. 1.8e-279;
XX Matches 1092; Conservative 25; Mismatches 67; Indels 15; Gaps 7;
Qy 387 MEAHOEVVAEVHVSTVEKTEEBEGGGEAGGVVVEGTGSLPPEKLAEPQVEPQEAEP 446
Db 1 MEAHOEVVAEVHVSTVEKTEEBEGGGEAGGVVVEGTGSLPPEKLAEPQVEPQEAEP 60
Qy 447 EELMKSRMCVSGGDHTQLTDLSPPEKTLPKHPEGIVSEVEMLSQERIKVQGSPLKLF 506
Db 61 EELMKSRMCVSGGDHTQLTDLSPPEKTLPKHPEGIVSEVEMLSQERIKVQGSPLKLF 120
Qy 507 SSSGLKKLGGKKGKGGGDEEPGEYQIHITESPESADEQKGESASSPEEPTTCL 566
Db 121 SSSGLKKLGGKKGKGGGDEEPGEYQIHITESPESADEQKGESASSPEEPTTCL 180
Qy 567 EKGPLEAPODGEAEEGTTSDGEKKRE---GITPWASFKKMVTPKKVRPSPESDKEEL 622
Db 181 EKGPLEAPRMGKLRKELLR-GEKKRKDHSLGI-----LQKGDGTQETVRRPSSDKEEL 234
Qy 623 EKVKSATLSSTDSTVSEMDEVTGVEEQPEPKRVDTSVSWEALICVSSSKRKARKA 682
Db 235 EKVKSATLSSTDSTVSEMDEVTGVEEQPEPKRVDTSVSWEALICVSSSKRKARKA 294
Qy 683 SSSDDEGGPRTLGGDSHRAEASKDXEAGTDVAPASTQEQDAQGSSSPSPAGSPSEGG 742
Db 295 SSSDIR-GRPRTLGGGQSRRGQRRSRDAPASTQEQDAQGSSSPSPAGSPSEGG 353
Qy 743 VSTWESFKRLVTPRKSKSKLEEK-AEDSSVEQLSTIEPSEESWVS--IKKPIGRRK 799
Db 354 VSTWESFKRLVTPRKSKSKLEEKAGRTLIVGAGCPLRSNRVKNLGFLLRNSSPDGR 413
Qy 800 KRADGKQEQATVEDSGPVEINEDDPNVPAVVPLSEYNAVEREKMAEQNTLPOLLCAVY 859
Db 414 KQMGREQEQATVEDSGPVEINEDEPDVPAVVPLSEYDAVEREKMAEQNALSCWGCV- 472
Qy 860 VSELSKTLVHTVSVAVIDCTRAVTSVEERSPSWISASVTEPLEHTAGEAMPVVEVTEK 919

```


Db 473 VSEBISKLTVHTVAVAVIDGTRAKTSKEERSBWSIASVTEPLEHTAGEMPVEVTEK 532
Qy 920 DIIAETPVLTVLPBEGDAHDMVTSEVDTSEAVTAETSEALRTEETVETASGAETTT 979
Db 533 DIIAETPVLTVLPBEGDAHDMVTSEVDTSEAVTAETSEALRTEETVETASGAETTT 592
Qy 980 DMVSAVSGTSPDTTEATPVQEVESGVLDTREEROTQILQAVAKVVEESQVPATQ 1039
Db 593 DMVSAVSGTSPDTTEATPVQEVESGVLDTREEROTQILQAVAKVVEESQVPATQ 652
Qy 1040 TVORTGSALKEVEEVEDSEVLASEKEDVMPKGPVQEGAEHLAQSEGTQATPEELE 1099
Db 653 TVORTGSALKEVEEVEDSEVLASEKEDVMPKGPVQEGAEHLAQSEGTQATPEELE 712
Qy 1100 VPEVTADVAVATCOVILKIQOLMEQAVAPESSETLTDSETNGSTPLADSDTADGTQDET 1159
Db 713 VPEVTADVAVATCOVILKIQOLMEQAVAPESSETLTDSETNGSTPLADSDTADGTQDET 772
Qy 1160 IDSQDSKATAAVRQSVTEBEAATQKEEPTLPPNNVAQEHGEEPRDVLPTQOELT 1219
Db 773 IDSQDSKATAAVRQSVTEBEAATQKEEPTLPPNNVAQEHGEEPRDVLPTQOELT 832
Qy 1220 AAADVVLAKTEVGEGEVDMLDGEKVEKEOEYFVHSGPNSOKADVTYDSEVMGVAQOE 1279
Db 833 AAADVVLAKTEVGEGEVDMLDGEKVEKEOEYFVHSGPNSOKADVTYDSEVMGVAQOE 892
Qy 1280 KESTEVQSLSEEGEMETDVEKEKRETKPEQVSEBEOETAPPEHEGTGYKPEVLTLDMP 1339
Db 893 KESTEVQSLSEEGEMETDVEKEKRETKPEQVSEBEOETAPPEHEGTGYKPEVLTLDMP 952
Qy 1340 SERGALSLGSGPSLPDODKAGCIEVQVQSLDTVTQTAENVEKVIETVVISSEGEPE 1399
Db 953 SERGALSLGSGPSLPDODKAGCIEVQVQSLDTVTQTAENVEKVIETVVISSEGEPE 1012
Qy 1400 CVGAHLPAEKSSATGHWTLQHAEDTVPLGPESQAESIPITVPAPSTLHPDLQGEIS 1459
Db 1013 CVGAHLPAEKSSATGHWTLQHAEDTVPLGPESQAESIPITVPAPSTLHPDLQGEIS 1072
Qy 1460 ASQRESEEDKPDAGPADGKSTALEKVLKABEILLEESKSNKIYLVNTQTAVDQFA 1519
Db 1073 ASQRESEEDKPDAGPADGKSTALEKVLKABEILLEESKSNKIYLVNTQTAVDQFA 1132
Qy 1520 RTEIPEHTAHYDSTQVPAACRLDSREPNRCWTMKMDAKMKHPVQPRBDLQVLTLEAM 1578
Db 1133 RTEIPEHTAHYDSTQVPAACRLDSREPNRCWTMKMDAKMKHPVQPRBDLQVLTLEAM 1191

RESULT 3
ID ABB97448 standard; Protein, 1783 AA.
AC ABB97448;
XX 27-JUN-2002 (first entry)
XX
XX
XX Novel human protein SEQ ID NO: 716.
XX
XX Human; anti-inflammatory; anti-inflammatory; immunomodulator;
XX anti-inflammatory; cerebroprotective; cytoprotective; rheumatic; gene therapy;
XX neuroprotective; antiparkinsonian; protein therapy; EST;
XX expressed sequence tag.
XX
XX Homo sapiens.
XX
XX WO200222660-A2.
XX
XX 21-MAR-2002.
XX
XX 10-SEP-2001; 2001WO-US26015.
XX
XX 11-SEP-2000; 2000US-0659671.
XX
XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
PI xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX
XX WPI; 2002-292408/33.
DR N-PSDB; ABN32634.
XX
PT An isolated polynucleotide for treating diseases associated with its
PT encoded polypeptide such as cancer and multiple sclerosis -
XX
XX
PS Claim 20; SEQ ID NO 716; 509pp; English.
XX
XX The present invention provides the protein and coding sequences of 444
XX novel human proteins. These were isolated from expressed sequences tags
CC (ESTs). They can be used to stimulate cell growth, to regulate
CC hematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
CC e.g. in burn treatment, to regulate the immune system e.g. to treat
CC multiple sclerosis, to regulate activin or inhibin e.g. to treat
CC infertility, to regulate haemostasis or thrombolysis e.g. to treat
CC stroke and cancer, to screen for drugs, to treat inflammatory conditions
CC e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.
CC Parkinson's disease. The present sequence is a protein of the invention.
XX
SQ Sequence 1783 AA;
Query Match 53.7%; Score 4334; DB 23; Length 1783;
Best Local Similarity 58.2%; Pred. No. 5,6e-223;
Matches 962; Conservative 188; Mismatches 401; Indels 102; Gaps 32;
Qy 1 MGAGSTTEORSPEO-PAGSDPTSEVLVSGHPAAEAS-GAAGDPADA--DPATLPPKNG 56
Db 1 MGAGSTTEORSPEO-PAGSDPTSEVLVSGHPAAEAS-GAAGDPADA--DPATLPPKNG 60
Qy 57 QLSVNGVABGDVHVEENQEQ-----EEVDEVDGQRESDVREKDR 102
Db 61 QLSVNGVABGDVHVEENQEQ-----EEVDEVDGQRESDVREKDR 120
Qy 103 VEEAANSTAVEDITKQGEETSEITIQIPASENNVEMQOPASQANDVGFKVFPRVG 162
Db 121 DKEAATSAVAVHDITDQGEETPEIIPSSSENLBELTQPPESQANDVGFKVFPRVG 180
Qy 163 FKFTVKKDKNEKSDPTVOLLTVKQDEGEAEASVAGGHQSPSVTAVGESKSEKOS 222
Db 181 FKFTVKKDKNEKSDPTVOLLTVKQDEGEAEASVAGGHQSPSVTAVGESKSEKOS 235
Qy 223 TEKQEGTLKQESSTEIPLQAESDQAAEEBAKQDEGEKQEKPTSPSPSPVNSSETTS 282
Db 236 TEKQEGTLKQESSTEIPLQAESDQAAEEBAKQDEGEKQEKPTSPSPSPVNSSETTS 294
Qy 283 SPKPFTHGMAGMKKTSPFKKSKEDDLFTAKRKEQEAEXYDEBEKTEBPASEE---- 337
Db 295 TPKKFPTQGMAGMRKKTSPFKPKEDVEASEKKKEQPEKQTEEDDEKAEVASEKLTASE 354
Qy 338 ----QEPADTDQARLSADYKVELPLEDOVDLEASSEKCAPLATEVPEKMEAHQ-E 392
Db 355 QAPQOEPRESAHEPRLSAHEKVELPSEBQVSGQSPSEKAPALATEVPEKMEAHQ-E 414
Qy 393 VVAEHNSTVEKTEEEGGGGEAEAGVVEGTGSLPPEKLAEPQEVQEAPEELMK 452
Db 415 VVAEHNSTVEKTEEEGGGGEAEAGVVEGTGSLPPEKLAEPQEVQEAPEELMK 466
Qy 453 REMCVSGGDHDTQLDLSPEEKLTPKHPBGIVSEVEMLSQERIVQSPPLKLFSSGGLK 512
Db 467 KETCVSGEDPTQADLSPDEKVLKSPPEGVYSEVEMLSQERIVQSPPLKLFSSGGLK 526
Qy 513 KLSGKKQKQKGGGDEPGEVYQHHTSPESADQKGEASASSPEPEETTCLEKGL 572
Db 527 KLSGKKQKQKGGGDEPGEVYQHHTSPESADQKGEASASSPEPEETTCLEKGL 585
Qy 573 APQGEAEETGTSDEGKKRREGITPMASFKKMVTPKKVRPSESDEKEELEKVASATLS 632
Db 586 VQDGEAEAGATSDGKKRREGITPMASFKKMVTPKKVRPSESDEKEELEKVASATLS 645


```

Db 235 TEKEBETLKREQSAHEISPPAEBSGOAV--EECKEBSGEKQEKPEKPSKASAPTSPTVTSNGS 293
Qy 283 SFKKFFTHGMAWKRKTSPFKSKEDDLETAERKEQEAKEYDEBEKEKTEBPASBE----- 337
Db 294 TFKKFFTHGMAWKRKTSPFKSKEDDLETAERKEQEAKEYDEBEKEKTEBPASBE----- 353
Qy 338 ----QEPABEDTDQARLADYKVELPLEDQVGDLEASSEKCAPLATVPPREKMAHQ-E 392
Db 354 QAHPOBPASAHBEPRLSAEYKVELPSEBOVSGSGPSEKRAPAPLATEVPBEKLEVHQEE 413
Qy 393 VVAHVHSTVKTETEEBOGGAEAGVVEGTESELPREKLAEPQEVQOAPAEPLMKLS 452
Db 414 VVAHVHSTVKTETEEBOGGAEAGVVEGTESELPREKLAEPQEVQOAPAEPLMKLS 465
Qy 453 REMCVSGGDHTQTLTDLSPBEKTLPPHREGIVSEVAMLSQERIKVQSGPLKLFSSGLK 512
Db 466 KETCVSGEDPTQAGADLSPEKVLSPREGVSEVEMLSQERIMKQVQSLKLFSTGLK 525
Qy 513 KLSGKKQKRGKGGGDEPGEVQHHTSPSADQKESASSSPEBEETTCLEKPLE 572
Db 526 KLSGKKQKRGK--GGDBESGHTQVPADSPSQEQKESASSSPEBEETTCLEKGLAE 584
Qy 573 APQGEAEGETTSGEKKREGITTPWASFKKMYTPKKRVRRPSESDKEBELEKVKSATLS 632
Db 585 VQOQGEAEGETTSGEKKREGITTPWASFKKMYTPKKRVRRPSESDKEBELEKVKSATLS 644
Qy 633 TDSVSEMODEVKTVGEBOKEPEPKRRVDTSVSWEALICVSGSKKRRARAKASSDDEGGPR 692
Db 645 TESTASEMOEEMKGSVEBPKEPEPKRRVDTSVSWEALICVSGSKKRRARAKASSDDEGGPR 704
Qy 693 TLGGDSHRAEASAKXKAGTDAVPASTOBODQAGSSSPERAGSSEBGVSTWMSFKRL 752
Db 705 AMGGHQAQADENAGKXKKEGTGDIILAGSQHDPGQSSSPERAGSSEBGVSTWMSFKRL 764
Qy 753 VTPRKKSXSKLEKXKAD-----SSVQLSTBELPSEBSVSIKKFIPGRKKRAGAKXQEO 808
Db 765 VTPRKKSXSKLEKXKEDSIAGSGVHSTPDTEPKESVWSIKKFI PGRKKRAGAKXQEO 824
Qy 809 ATVEDSGVEINEDDPNPAVPAVPLSEYNAVEREKME---AQNTLPLOLLGAVVSEELS 865
Db 825 APVEBAGTGANEDSDVPAVPAVPLSEYDAVEREKMEAOQAQAGAQPEQKATEVSKELS 884
Qy 866 KTLVHTSVAVIADTGRAVTSVEERSPSWISASVTPLEHTAGEMPVPEVTEKIIA-E 924
Db 885 ESQVHMAAAVADGTTRAFTIIEERSPSWISASVTPLEHTAGEMPVPEVTEKIIA-E 944
Qy 925 ETPVLTQTLPEGKXADHDMVYSEVDTSEAVATETSEALTREEVTEBASGAETTDWYSA 984
Db 945 EPPVTEPLPENREARGDTTVSEALTREAVTAETTAGPLGSEEGTESABEETTEMVSA 1004
Qy 985 VSQLTDSPTTEEATPVQEVESGVLDTEEBERQTOALLQAVADKVEESQVPAE----- 1038
Db 1005 VSQLTDSPTTEEATPVQEVESGVLDTEEBERQTOALLQAVADKVEESQVPAE----- 1064
Qy 1039 -QTVORTSKALEKEYEVBEDSEVLASEKXKVMKPGVQEAHAHLAOGSTGATPES 1097
Db 1065 LQPVQRA-----EAERBEBOQEAAGLKKETDVLKVDPAQEAKEKTPFGKVVGGTTES 1118
Qy 1098 LE-VBEVYADVDH---VATTCV-----IKLQOL-MEOAVAPESSELTLDSENGSTPLAD 1147
Db 1119 FEKAOQVTESSLESELVTTQOETLAGVSOEMWEOALPPPSVETPTDSETDGSIPTVAD 1178
Qy 1148 SPTADGTQODETITDSQSKATPAVAGSQVTEEBEATAQKEBESSTPNNVPAOEHEGEBFG 1207
Db 1179 FDPARTTQKDEIVEIHEBEVNASGTQSGGTEAENPAQKERRPA-PSSVPEBEETKEQSK 1237
Qy 1208 -RDVLEPTQOELITAAVAVPLAKTEVQGBEVDMLDGEKVK-----EBOEVFVHSG--PNS 1259
Db 1238 MEDTEHHTDKVSVETVSLSTKEGNO--EADQVADKTKVDVPFEGLEGSIDTGITVSR 1295
Qy 1260 QKADVYTTDSEVGVAGCOEKESTEVQ-----LSLEBEMTVDVKEKRETKPEOVSE 1314
Db 1296 EKVTEVALKGBETBAECKDQDALELQSHAKSPSPVEREMVVOVEREKTEAPFTHVNEE 1355

```

```

Qy 1315 G-EQETAPHEGTYGKPVLTLMPSSEERKALGSLGS--PSLPDQKACIEVQVOSLD 1372
Db 1356 KLEHETVTVTSEEVSKQLQTVNVPILIDGAKEVSSLEGSPPCLQGBEAVCTKIQVOSSE 1415
Qy 1373 TTVTQTAEAVEKVI--ETVWISETGESPECVGAHILPAEKSSATGHWTLQHAEDVPLG 1430
Db 1416 ASFTLTAAEBEKVLGTANILETGETLEPAGAHVLBEKSSSEKNEDEPAAHPGEDAVPTG 1475
Qy 1431 PESQAESIPITVPAPRESTLHPDLGEISASQERSEEDKPDAGPDADGKESTAIKVL 1490
Db 1476 PDCQAKSTPVISATTKGLSSDLEGEKTSILKWKSDPEVDQVACQEV--KVSVALIDL- 1532
Qy 1491 KAEPB--ILLESNSKIVANVLOTAVDOFART-ETAPETHANDSQOVACRLDSEEPN 1547
Db 1533 --EPENGILELETSSKLONIIOTAVDOFRTBEETATEMLTSELQTOAHVIRKADSODAG 1590
Qy 1548 RCWTK-----MKDAKMHPVQPREDLQVLTVEA 1577
Db 1591 QETKEGEBFOASADETPTITSKESESESTAVGQA 1625

```

```

RESULT 5
AAM53863
ID AAM53863 standard; peptide; 1780 AA.
XX
AC AAM53863;
XX
DT 13-JUL-1998 (first entry)
XX
DE Human gravin polypeptide.
XX
KW Gravin; kinase anchoring protein; type II regulatory subunit; PKA; PKC;
KW cAMP-dependent protein kinase; protein kinase C; autoimmune disease;
KW Myasthenia gravis; nicotinic acetylcholine receptor.
XX
OS Homo sapiens.
XX
PN US5741890-A.
XX
PD 21-APR-1998.
XX
PF 19-DEC-1996; 96US-0769309.
XX
PR 19-DEC-1996; 96US-0769309.
XX
PA (UYOR-) UNIV OREGON HEALTH SCI.
XX
PI Klauch TM, Nauert JB, Scott JD,
XX
XX WPI; 1998-260552/23.
XX
DR N-PSDB; AAV23545.
XX
PT New polypeptide fragments of protein kinase binding protein gravin -
PT are useful for the study of modulation of action between gravin and
PT protein kinase (s)
XX
PS Example 1; Column 19-32; 32pp; English.
XX
SS This sequence corresponds to the human gravin polypeptide, and represents
CC a polypeptide of the invention. The polypeptides are fragments capable of
CC binding to type II regulatory subunit of cAMP-dependent protein kinase
CC (PKA). Gravin is a kinase anchoring protein that binds to type II
CC regulatory subunits of PKA and protein kinase C (PKC). Gravin is also an
CC antigen of the autoimmune disease Myasthenia gravis (MG), where a patient
CC develops antibodies against their own nicotinic acetylcholine receptors.
CC The polypeptides are useful for providing analogues of gravin in the
CC study of the modulation (e.g. blocking, inhibiting and stimulating) of
CC interactions between gravin and kinase. The peptides are involved in the
CC modulation of gravin-kinase interactions.
XX
SQ Sequence 1780 AA;

```

Query Match	52.9%	Score 4274.5	DB 19	Length 1780
Best Local Similarity	57.8%	Pred. No. 8.5e-220		
Matches 957	Conservative 187	Mismatches 402	Indels 109	Gaps 36
Qy	1	MGAGSSTEQRSPEQ--PAGSDTPELSVLGHGPAARAS-GAAGDPADA--DPATKLPOKNG	56	
Db	1	MGAGSSTEQRSPEQPPGSSTPAEPEPGCGPSAAGPDTTADPAIAASDPATKLQKNG	60	
Qy	57	QLSSVNGVAEQGDVHVQBEHQEQ-----EEEVVDVDEGVQRESDVREKDR	102	
Db	61	QLSTINGVAEQDELSQLQEGDLNGKALNGQALNGQALNSQEEEEVIIVTEGVQDSEDSVERDS	120	
Qy	103	VEEMAANSTAVEDITYKQOQETSEIIEQIPASNNVEMVQPAESQANDVGPKVPKPVG	162	
Db	121	DKEMATKSAVVHDIITDDQOENRN-IEQIPSSNESLEELTQPTESQANDIGPKVPKPVG	179	
Qy	163	FKFTVKDKNEKSDTVQLLTVKKDREGAEASVGAGDHOEPSVETAVGESASKESELKQS	222	
Db	180	FKFTVKDKTEKPDIVQLLTVKKDBEGA--AGAGDHQDPLS--GAGEAASESEPKQS	234	
Qy	223	TEKQEGTLKQBSSTEIPLQAESDQAABEEAKDEGEEKQEKFTKSPSPSPSPVNASSETTS	282	
Db	235	TEKPEETLKREQSHAEISPPAESQAV-EECKEEGEEKQEKPSKSAESPTSPVTSCTS	293	
Qy	283	SPKPFTHGWAGWRKKTSTFKSKSDKEDDLTAETAEKTKQAEKAVDDEBEKKEKTEPASEE----	337	
Db	294	TFKPFTHGWAGWRKKTSTFRPKPEDEVEASEKKQEPEKVDTEBDGKAEVASEKLTASE	353	
Qy	338	---QEPADETDQARLSADYKEVELPLEDQVCDLNASSEKCAPLATEVDFDEKMAHQ-E	392	
Db	354	QAHPQEPAESAHPEKLSABYKEVELPSEBQVSGSGQSPSEKPAPLATEVDFDKIEVHQEE	413	
Qy	393	VVAEVHVSTVKEETEEQCGGGAEGGVVVGTEGSLPPEKLAEPQEVPOEAPPAEELMKKS	452	
Db	414	VVAEVHVSTVEETEEQ-----KTEVEETAGSVPAEELVGMDAEPOEAPPAKELVKL	465	
Qy	453	REMCVSGGDHQLTDLSPPEKTLPKHPGPIVSEVEMLSQERIKVQGSPLKKLPSSGLK	512	
Db	466	KETCVSGDPTQCADLSDPEKVLSPKPPGVVSEVEMLSQERMKVQGSPLKKLPSTGLK	525	
Qy	513	KLSGKKQKRGKGGDEBPGEYQHHTSPESADQKSGSSASSPEEPEETCCLKGPLE	572	
Db	526	KLSGKKQKGRK-GGGDEBSGHTQVAPDSPDSQEQSGSSASSPEEPEETCCLKGLAE	584	
Qy	573	APQGEABEGTTSQGEKKREGITPWASPKMVTPKKRVRRPSESDEKELEKVKKSATLSS	632	
Db	585	VQQGEABEGATSDGEKKREGVTWASPKMVTPKKRVRRPSESDEKDELDKVKKSATLSS	644	
Qy	633	TDSTVSEMQDEVKTVGEBQPEPKRRVDTSVSWEALICVGSKKRKARAKSSDDGGPR	692	
Db	645	TESTASEMOEBMKGSEBPPEPKPRKVDTSVWEALICVGSKKRRRRSSSDEGGPK	704	
Qy	693	TLGGDSHRAEASDKKAGTADVAPASTQEQDAQGSSSPAPAGSPSEGEVSTWESFKRL	752	
Db	705	AMGDDHQKADAGDKKETGTGILAGSQEHDPGQSSSPQEQAGSPTEGEVSTWESFKRL	764	
Qy	753	VTPRKKSCLKEEKAED---SSVEQLSTETSPSEESWVSTKKPIGRRKKRAGKQEQ	808	
Db	765	VTPRKKSCLKEEKSEDSIAGSGVBSHTDTPTEPGKESWVSTKKPIGRRKKRPPGKQEQ	824	
Qy	809	ATVEDSGPEINEDDPNYPVAVPLSEYNAREKME---AQNTLPLQLLGAIVVSEELS	865	
Db	825	APVEDAGPTGANEDSDVPVAVPLSEYDAVEREKVMAQQAQKAGQEPQKAATEVSEKLS	884	
Qy	866	KTLVHTVSVAVIDGTRAVTSVEERSPSWISASVTBPLEHTAGEAMPVVEVTEKDIIA-E	924	
Db	885	ESQVHMAAAVADGTAAATIIERSPSWISASVTBPLEQVEAEAAALLTEEVLEREVIABE	944	
Qy	925	ETPVLTOTLPEGKAHDHDMVTSEVDFTSQVATATSTSEALRTEEVTEASGAEETDMVSA	984	
Db	945	EPPIVTEPLPNNRARGTQVSEAEITPEAVTAAETAGPLGSEEGTEASAAABETTMVSA	1004	
Qy	985	VSQUTDSPDTTEATPPQVEBSGVLDTEEBERQTOAILQAVADVKKESQVPAAT-----	1038	

PR 19-DEC-1996; 96US-0769309.
 XX (UYOR-) UNIV OREGON HEALTH SCI.
 PA
 XX KLauck TM, Scott JD, Nauert JB;
 XX WPI; 2000-523763/47.
 DR N-PSDB; AAA74903.
 PT Novel polynucleotides useful for detecting gravin in patients suffering
 PT from Myasthenia gravis encodes cAMP-dependent protein kinase-binding
 PT polypeptide and protein kinase C-binding polypeptide of gravin -
 XX
 XX Claim 1; Column 35-45; 34pp; English.
 CC The present sequence is the protein sequence of human gravin. Gravlin is
 CC an A-kinase anchoring protein (AKAP) which is involved in the
 CC localisation of cAMP dependent protein kinase A (PKA) via interactions
 CC between the RII binding region and the PKA regulatory subunit RII. Gravlin
 CC is also an antigen found in myasthenia gravis sufferers, and it is
 CC thought that antibodies to it may be useful in modulating the binding of
 CC PKA, and thus aid in the treatment of the disease. The gravin coding
 CC sequence was isolated by first screening a human umbilical vein
 CC endothelial cell cDNA library with serum from a myasthenia gravis
 CC patient, and then searching a human heart cDNA library for sequences
 CC resembling the isolated sequence. This was done because the first
 CC sequence obtained was shown to be shorter than the full length cDNA.
 XX
 SQ Sequence 1780 AA;
 Query Match 52.9%; Score 4274.5; DB 21; Length 1780;
 Best Local Similarity 57.8%; Pred. No. 8.5e-220;
 Matches 957; Conservative 187; Mismatches 402; Indels 109; Gaps 36;

QY 573 APDGEAEGETTSDGKKRRCITPMASFKKMTPKKRVRRPSSDKEELEKYSATLSS 632
 DB 585 VQDGEAEGETTSDGKKRRCITPMASFKKMTPKKRVRRPSSDKEELEKYSATLSS 644
 QY 633 TDSVSEMODEVKTGVEEQKPEEPKRRVDTSVSMALICVSSSKKARKASSSDDEGGPR 692
 DB 645 TESTASMODEEMKGVSEPRPEEPKRRVDTSVSMALICVSSSKKARKASSSDDEGGPR 704
 QY 693 TLGDSHRAEASKKAGTDAVPASTQEDDQAGSSPEPPASSPSEGEVSTWESFKL 752
 DB 705 AMGDHQAEDAGKXETGTGDIAGQEHDPGQSSSPQOASSPTEGEGVSTWESFKL 764
 QY 753 VTPRKKSKLSEKAE-----SSVEQLSTEIPSRSESWSIKKFIPGRKKADGQEQ 808
 DB 765 VTPRKKSKLSEKSEDSINGSGVESHPTPEPKESWSIKKFIPIGRKKRPDQEQ 824
 QY 809 ATVEDSGPVEINEDDPVAVPVPVLESEYNAVERKME---AQNTLELPOLLGAVVSELS 865
 DB 825 APVDAFTGANEEDSDVPVAVPVPVLESEYNAVERKMEQAQKGAEOEQKAAPEVSKELS 884
 QY 866 KTLVHTVSAVIDGTAVTSVEERSPSWISASVTEPLEHTAGEMAPVEYTEKDIIA-E 924
 DB 885 ESQVHMAAAVAADGTAAITIIERSPSWISASVTEPLEQVEAEALITTEVLEREVIAEE 944
 QY 925 ETPVLQTLPEGRKANDMTSEVDFTESEAVTATESEALRTEETASGAEETDMVSA 984
 DB 945 EPTVTEPLPNNRRARGDTVSESEALTEPAVTAETAGLPLOSEGTASAEETTEMVSA 1004
 QY 985 VSQLTSDPTTEATPVQVEESGVLTDEEERQTOAILQAVADVKEESQVPAT----- 1038
 DB 1005 VSQLTSDPTTEATPVQVEESGVLTDEEERQTOAILQAVADVKEESQVPAT----- 1064
 QY 1039 -QTVQRTGSKALEKVEEVEEDSEVLASBKEDVMPKGPVQEAAGHMLAQSETOQAPPS 1097
 DB 1065 LQPVQRA-----BAERPEEQAEASGLKETDVVLKXDAQAKTEPTQGVVQQTTPES 1118
 QY 1098 LE-VPEVADVDH---VATQGV-----IKLQOL-MEOAVASESETLTDSENGSTPLAD 1147
 DB 1119 FEKAPQVTESESELVTTCAETLAGVQSEMEQOAIIPDSVETPTDSTEDSTPAD 1178
 QY 1148 SPTADGTQOODETIDSDSKATAAVROSQVTEEAATQKEPSTLPNNVPAQOEHHGEPG 1207
 DB 1179 FPAETGTOQDEIVELHEENENHIV-PVKGTEAEAVPAQKEPRA-PSSFVQOETKEQSK 1236
 QY 1208 -RDVLEPTQOELTAAVAVLAKTEVQGEVDMLDEKVK-----EBOEVFVHSG--PNS 1259
 DB 1237 MEDTLEHTDKESVETVAILSTKEGTQ--EADQYADEKTKDVPFPEGLEGSIDTGITVSR 1294
 QY 1260 QKADVTYTDSEVMGVAQOEKSESTEVQ-----LSLEGEMETVYEKERTKQVSEE 1314
 DB 1295 EKVTEVALKSGTEBAEKQKDALTELQSHAKSPSPVREAVVQVERKTEAEAPHVVEE 1354
 QY 1315 G-EQETAPREHGTGKGVLTLDMPSSERKALGSLGS-PSLPDQDAGCIEVQVSLD 1372
 DB 1355 KLEHTATVTVSEEVSKQLQTVNVPITIDGAKESVLSBSPPCGQGEAVCTKIQVQSS 1414
 QY 1373 TTVTQTAAEVKVI--ETVVISETGESEPCVGAHLLPAEKSSATCGHWTLOHAETVPLG 1430
 DB 1415 ASFTLTAABEEKVLGETANILETGETLEPAGAHVLTLEKSESEKNEDEAHPGAEAVPTG 1474
 QY 1431 PESQAESPIITYTPARESTLHDLQELISASORESESEEDKPDADPDGKSTAIKYL 1490
 DB 1475 PDCQAKSTFVIVSATTKGLSDDLGEKTTSLKMSDEYDEQVACQEV--KVSVAIEDL- 1531
 QY 1491 KAEPE--ILELESKNKIYLVNITQAVDQFART-ETAEETHAVDSQOTQVPAICRLDSREBN 1547
 DB 1532 --EPENGILEFTKSKLVQNIQITAVDQFARTETAEETHAVDSQOTQVPAICRLDSREBN 1589
 QY 1548 RCMWK-----MKDAKMKHPVQPREDDQVLTVLEA 1577
 DB 1590 QETEKGEPEPQASQADEPTITSKESESTAVGOA 1624

QY 1148 SDTADTQOQDETISQDSKATPAVQSOVTEFEATTAQKEBSSTIPNNVPAQEHGEBPG 1207
 Db 1188 FDAPGTQKDEIVEIHEENENAVSGTQSGTEAAVPAKERPBA-PSSRFVQETKEQSK 1246
 QY 1208 -RDVLEPQOGLTAAVAVLAKTEVGOGEVMDLGEKYK-----EEOGVFVHSG--PNS 1259
 Db 1247 MEDTLEHTDKVESVETVSTLSKTGTO--EADQVADKTKVDPFEGLEGSTIDTGTISR 1304
 QY 1260 QKADVTYDSEVMGVAGCOEKESTEVQS-----LSLEGEMTDVKEKRETK-DEQVSE 1313
 Db 1305 EKVEITVALKGESTEEAECKKDALFLQSHAKSPSPVEREMVQVVEREBYQXQSPTHVNE 1364
 QY 1314 EG-EOETAPAEHEGTGKPVLTLDMPSSERKALGSLGSS-PSLDQDKAGCIEVQVQSL 1371
 Db 1365 EKLHEHTAVTVSEEVSKQLLTQVNVPIIDGAKESVLSLESPPCGQEBAVCTKIQVQSS 1424
 QY 1372 DTTVQGTAEAVKVI--ETVVVSEGTGESPECVGAHLPAEKSSATGSHWTLQHAEDTVPL 1429
 Db 1425 EASFLLTAAAEKVKVGETANILETGETLEPAGAHVLEKSEKKNEDPAHPGEDAVPT 1484
 QY 1430 GSESQAESIPITVTPAPESTLHPDLQGEISASQERSEBEDKPDAGPDAGKESPAIEKY 1489
 Db 1465 GPDCAKSPVIVSATTKGSLSDLEGEKTSLKMKSDVDVQVACQEV--KVSVAIEBD 1542
 QY 1490 LKAPELLLESKSNKIVLNTQTAVDQFART-ETAPETHAVDSQTVQVACRLDSREPMK 1548
 Db 1543 LEPENGILLEFTKSKVLQNIITQTAVDQFVTEETATEMLTSELQTAHVIRADSDAQ 1602
 QY 1549 CWTK-----MKDAKMKHPPQPRREDQVLTVLEA 1577
 Db 1603 ETEKEGEDQASADQETPTTSAKESESTAVGQA 1636

RESULT 8

AA053862
 ID AA053862 standard; peptide; 292 AA.

AA053862;

13-JUL-1998 (first entry)

Gravin polypeptide fragment, residues 265-556.

Gravin; kinase anchoring protein; type II regulatory subunit; PKA; PKC;
 CAMP-dependent protein kinase; protein kinase C; autoimmune disease;
 Myasthenia gravis; nicotinic acetylcholine receptor.

Homo sapiens.

US5741890-A.

21-APR-1998.

19-DEC-1996; 96US-0769309.

19-DEC-1996; 96US-0769309.

(UYOR-) UNIV OREGON HEALTH SCI.

Klauck TM, Nauert JB, Scott JD;

WPI; 1998-260552/23.

New polypeptide fragments of protein kinase binding protein gravin -
 are useful for the study of modulation of action between gravin and
 protein kinase(s)

Claim 3; Column 17-20; 32pp; English.

This sequence corresponds to residues 265-556 of gravin, and represents
 a polypeptide of the invention. The polypeptides are fragments capable of
 binding to type II regulatory subunit of CAMP-dependent protein kinase

(PKA). Gravin is a kinase anchoring protein that binds to type II
 regulatory subunits of PKA and protein kinase C (PKC). Gravin is also an
 antigen of the autoimmune disease Myasthenia gravis (MG), where a patient
 develops antibodies against their own nicotinic acetylcholine receptors.
 The polypeptides are useful for providing analogues of gravin in the
 study of the modulation (e.g. blocking, inhibiting and stimulating) of
 interactions between gravin and kinase. The peptides are involved in the
 modulation of gravin-kinase interactions.

Sequence 292 AA;

Query Match 11.8%; Score 952.5; DB 19; Length 292;

Best Local Similarity 66.1%; Pred. No. 2,7e-43;

Matches 199; Conservative 30; Mismatches 53; Indels 19; Gaps 4;

QY 254 KQEGEKEKPTSPSPSPVNSSETTSRKFFTGAMGRKKTSPKSKEDDLTAE 313
 Db 1 KEEGEEKKEKPSKASAPSPVTSSTGTFKFFTGAMGRKKTSPKSKEDDLTAE 60
 QY 314 KRKQAEKVDSEKKEKTEPASE-----QEPADTDQARTLSADYKVELPLEDOV 364
 Db 61 KKKQEPKVTBEDGKAEVASEKLTASQAHPOEPASAEHRLSAHYKVELPSEQV 120
 QY 365 GDLEASSEKCAPLATEVFEDEKMEAHQ-EVVAEVAHVSTVEKTEBEEGGGGEABGVVVG 423
 Db 121 SSGQSPSEKAPLATEVFEDEKIEVHOEVAVAHVSTVEERTSEQ-----KTEVEE 172
 QY 424 TGESIPPEKLAEPQEVPOEAEPAEELMKSRMCVSGDHTQLTDSPEKTLPHKPGIV 483
 Db 173 TAGSVPAELVGMDAEPQEAEPAEELVLTAKETCVSGEDPTQGADLSPEKLSKPPGEGV 232
 QY 484 SEVENTSSOERIKYOGSPKLTLPSSGLKXSGKKQGGGGDEEPGEQHTHSPE 543
 Db 233 SEVENTSSOERIKYOGSPKLTLPSSGLKXSGKKQGGGGDEEPGEQHTHSPE 291
 QY 544 S 544
 Db 292 S 292

RESULT 9

AA015379
 ID AA015379 standard; protein; 292 AA.

AA015379;

26-JAN-2001 (first entry)

Recombinant human gravin polypeptide fragment #1.

Human; gravin; PKA RII binding site; myasthenia gravis;
 kinase anchoring protein; CAMP dependent protein kinase.

Homo sapiens.

US6090929-A.

18-JUL-2000.

19-DEC-1997; 97US-0994570.

19-DEC-1996; 96US-0769309.

(UYOR-) UNIV OREGON HEALTH SCI.

Klauck TM, Scott JD, Nauert JB;

WPI; 2000-523763/47.

Novel polynucleotides useful for detecting gravin in patients suffering
 from Myasthenia gravis encodes CAMP-dependent protein kinase-binding
 polypeptide and protein kinase C-binding polypeptide of gravin -

PS Example 4; Column 21-24; 34pp; English.

XX The present sequence is a recombinant version of the human gravin

CC protein. Gravin is an A-kinase anchoring protein (AKAP) which is involved

CC in the localisation of cAMP dependent protein kinase A (PKA) via

CC interactions between the RII binding region and the PKA regulatory

CC subunit RII. Gravin is also an antigen found in myasthenia gravis

CC sufferers, and it is thought that antibodies to it may be useful in

CC modulating the binding of PKA, and thus aid in the treatment of the

CC disease. This sequence was used in an assay to determine the ability of

CC gravin to bind protein kinase C.

XX Sequence 292 AA;

Query Match 11.8%; Score 952.5; DB 21; Length 292;

Best Local Similarity 66.1%; Pred. No. 2.7e-43;

Matches 199; Conservative 30; Mismatches 53; Indels 19; Gaps 4;

QY 254 KGEERKQEKPTKSPSPVNSSETTSFKKFFTHGAGWRKKTSPKSKEDDLLETA 313

DB 1 KEGEEKQEKPSKSAESPTSPVTSSTGTFKFFTHGAGWRKKTSPKSKDEVEASE 60

QY 314 KRKEQEAKEKVDDEEKEKTEPASEE-----QEPADTDQARLSADYEKVELPLEDOV 364

DB 61 KKKEQEPKVDTEEDGRAEVAEKLTAHQAPQEPAPSAEHPRLSAEYKVELPSEQV 120

QY 365 GDLEASSEEKCAPLATVDFEKEMAHQ-EVVAHVHSTVTEKEEQGGGGEAGGVVVEG 423

DB 121 SGSGPSEKPAPLATEVDFEKEVHQEEVVAHVHSTVTEKEEQ-----KTEVEE 172

QY 424 TGESLPKLAEPQEPQEPAPAEALMKSREMCVSGGDHQLTDLSPKEETLPHKPGIV 493

DB 173 TAGSVPAELVGMDAEPQEPAPAEALMKSREMCVSGGDHQLTDLSPKEETLPHKPGIV 232

QY 484 SEVEMLSQERIKVQSGPLKLFSSGLKGLKQKQKGGGDEPPGVEYQHIHTESPE 543

DB 233 SEVEMLSQERIKVQSGPLKLFSSGLKGLKQKQKGGGDEPPGVEYQHIHTESPE 543

QY 544 S 544

DB 292 S 292

RESULT 10

AAW53875

ID AAW53875 standard; peptide; 651 AA.

AC AAW53875;

XX 13-JUL-1998 (first entry)

DE Gravin polypeptide fragment, residues 1130-1780.

XX Gravin; kinase anchoring protein; type II regulatory subunit; PKA; PKC;

KW cAMP-dependent protein kinase; protein kinase C; autoimmune disease;

KW Myasthenia gravis; nicotinic acetylcholine receptor.

OS Homo sapiens.

XX US5741890-A.

XX 21-APR-1998.

XX 19-DEC-1996; 96US-0769309.

XX 19-DEC-1996; 96US-0769309.

XX (UYOR-) UNIV OREGON HEALTH SCI.

XX Klauck TM, Nauert JB, Scott JD;

XX WPI; 1998-260552/23.

XX

PT New polypeptide fragments of protein kinase binding protein gravin -

PT are useful for the study of modulation of action between gravin and

PS protein kinase (s)

XX Example 4; Column 49-54; 32pp; English.

XX This sequence corresponds to residues 1130-1780 of gravin, and represents

CC a polypeptide of the invention. The polypeptides are fragments capable of

CC binding to type II regulatory subunit of cAMP-dependent protein kinase

CC (PKA). Gravin is a kinase anchoring protein that binds to type II

CC regulatory subunits of PKA and protein kinase C (PKC). Gravin is also an

CC antigen of the autoimmune disease Myasthenia gravis (MG), where a patient

CC develops antibodies against their own nicotinic acetylcholine receptors.

CC The polypeptides are useful for providing analogues of gravin in the

CC study of the modulation (e.g. blocking, inhibiting and stimulating) of

CC interactions between gravin and kinase. The peptides are involved in the

CC modulation of gravin-kinase interactions.

XX Sequence 651 AA;

QY 1110 VATCOV-----IKLOQL-MEQAVAPESSETLTDSETNGSTPLADSDTADGTQODEIDSQ 1163

DB 6 VTTQCAETLAGVQSBMVMEQAIIPDPSVETPTDSETDGTGTPVADFAPGTOKDEIVEIH 65

QY 1164 DSKATAAARQSOVSTEEAATAOEESTLPPNVPAQEEHGEPEP-RDVLPTQOELTAAA 1222

DB 66 ENEVELV-PVRGTEAENPAQKPPA-PSSFVFOEETKEQSKMEDTLEHTDKESVET 123

QY 1223 VPVLAKEVQGEQEVMDLGEKVK-----EEQEVFVHSG--PNSOKAADVTYDSEVMGVA 1275

DB 124 VSILSKTEGTQ--EADQYADEKTKDVPFFEGLEGSIDTGTIVSRKVTVALKGEETEA 181

QY 1276 GCOEKESTEVQS-----LSLEGEMTDVKEKRETKPQOVSEEG-EOETAAPHEGTYG 1329

DB 182 ECKKDDALELQSHAKSPSPFVEREMVQVEREKTEAEPTHVNEEKLHETAVTVSEVSK 241

QY 1330 KPVLTLDMPSSBERGKALGSLGGS-PSLPDQKAGCTEVQVSLDTTDTTAAEAVEKVI-- 1386

DB 242 QLLQTVNVPIIDGAKESVSLGSPPLCQGEAEVCTKIQQVSEASEFTLTAAAEKVLG 301

QY 1387 ETVVISSETGESPECVGAHLLPAEKSSATGCHWTLQHAEDTVPGLGPSQAESPIIIVTPAP 1446

DB 302 ETANILETGETLEPACAHVLEEKSEKNEDEFAAHGDEDAVPTGPDCAKSTPVIYSATT 361

QY 1447 ESTLHPDLQGEISASORERSEEDKPDAGPDADGKESSTAIEKVLKAEPE--ILELESKN 1504

DB 362 KKGSLSDLEGEKTTSLKWSDEVDEQVACQEV--KVSVAIEDL---EPENGILELETKSS 416

QY 1505 KIVLVNIQTAVQDFART-ETAPETHAYDSQTVPAACRLDSRENRCTWK----MKDAKM 1558

DB 417 KLVQNIQTAVQDFVTEETATMTSLTSELQTAHVIKADSQDQAGQTEKEGEEPOASAQD 476

QY 1559 KHEVPQPFREDLQVLTVLEA 1577

DB 477 ETPITSKEESESSTAVGQA 495

RESULT 11

AAAB18410

ID AAAB18410 standard; protein; 651 AA.

XX AAAB18410;

XX 26-JAN-2001 (first entry)

XX Human gravin PKA regulatory subunit RII binding site assay peptide #3.

XX Human; gravin; PKA RII binding site; myasthenia gravis;

KW kinase anchoring protein; cAMP dependent protein kinase.

```

XX OS Homo sapiens.
XX US6090929-A.
XX 18-JUL-2000.
XX 19-DEC-1997; 97US-0994570.
XX 19-DEC-1996; 96US-0769309.
XX (UYOR-) UNIV OREGON HEALTH SCI.
XX Klauck TM, Scott JD, Nauert JB;
XX WPI; 2000-523763/47.
XX Novel polynucleotides useful for detecting gravin in patients suffering
XX from myasthenia gravis encodes CAMP-dependent protein kinase-binding
XX polypeptide and protein kinase C-binding polypeptide of gravin -
XX Example 2; Column 53-58; 34pp; English.
XX The present sequence is part of the human gravin protein. Gravin is an
XX A-kinase anchoring protein (AKAP) which is involved in the localisation
XX of CAMP dependent protein kinase A (PKA) via interactions between the
XX RII binding region and the PKA regulatory subunit RII. Gravin is also an
XX antigen found in myasthenia gravis sufferers, and it is thought that
XX antibodies to it may be useful in modulating the binding of PKA, and thus
XX aid in the treatment of the disease. This sequence was used in an assay
XX to identify the precise position of the RII binding site within the
XX gravin protein.
XX Sequence 651 AA;
SQ
Query Match 8.1%; Score 656; DB 21; Length 651;
Best Local Similarity 38.3%; Pred. No. 6.1e-27;
Matches 191; Conservative 80; Mismatches 188; Indels 40; Gaps 17;
QY 1110 VATCOV-----FKLOOL-MEOAVAPESSETLTDSEFNGSTPLADSDPTADGTQODETIDSQ 1163
DB 6 VTTQAEITLAVGKSGQEMWEOALPSPSVETPTDSDSTPVPADFPACTQKDEIVEIH 65
QY 1164 DSKATAVAQSOVTEEAATAQKEEPTLPNNVPAQEEHGEERPG-RDVEPTQOEITAAA 1222
DB 66 EENEVHLV-PVAGTEAEVPAQKERPPA-PSSFVFOEETKQSKMMDLEHTDKESVET 123
QY 1223 VVVLAKTEVGQGEVDWLDGKVK-----EQQEVFVHSG--PNSOKAAVTVDSVMGVA 1275
DB 124 VSILSTETGTO--EADQVADKTKDVPFPEGLEGSTIDTIGITVSRKVTVALKGBSTEEA 181
QY 1276 GGOEKSTEVOS-----LSLEGEMETDVEKEKRETKPQVSEEG-EQETAAPEHGTYG 1329
DB 182 ECKDDALELOSHAKSPSPVEREMVQVERKTEAEPIHVAHEKLEHETAIVTSEVSK 241
QY 1330 KEVLTLDMPSSERKALGSLGGS-PSLPQDRAKCTIEVOVOSLDTTGTOTAAVEKVI-- 1386
DB 242 QLLQGTNVNPIIDGAKESVSLGSPPCLGEEAVCTKIQVQSEASFTLTAABEEKVLG 301
QY 1387 ETVVIVSETSESPCVGAHLIPAKSSATGCHMTLOHAEPTVPLGPSSQASPTIITPAP 1446
DB 302 ETANILLETGETLEPPAGAHVLVEKSSSEKNEDFAAHGEBAVPTGPDCAKSTPVIIVSAT 361
QY 1447 ESTLHNDLGGELISASQRESEEDKPDAGDADGKESTAIKVLKAEPE--ILEESKN 1504
DB 362 KGLSSLDGEGKTSLSKMSDEVBOVACOEV--KVSVAIEDU--EPENGILEETKKS 416
QY 1505 KIVLVNIQTAVDQFART-ETAPETHAYDSQTVPAICRLDSREPNRCMTK-----MKDAM 1558
DB 417 KIVONITQITAVDQFVTEETATETMLTSLQTOAHVITKADSDAGQGETEKGEPEQASAD 476
QY 1559 KHPVPQPRDLOVLTIVLEA 1577
: : : : :

```

```

DB 477 ETPTSAKESSESTAVGQA 495
RESULT 12
ABB68397
ID ABB68397 standard; Protein; 2768 AA.
XX
XX ABB68397;
AC
XX
XX 26-MAR-2002 (first entry)
DE Drosophila melanogaster polypeptide SEQ ID NO 31983.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical.
XX Drosophila melanogaster.
XX WO200171042-A2.
XX 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US09231.
XX 23-MAR-2000; 2000US-191637P.
XX 11-JUL-2000; 2000US-0614150.
XX (PEKE ) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers EW,
XX WPI; 2001-656860/75.
XX N-PSDB; ABL12500.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX Disclosure; SEQ ID NO 31983; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins
XX (ABB57737-ABB72072).
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 2768 AA;
SQ
Query Match 8.0%; Score 642; DB 22; Length 2768;
Best Local Similarity 22.2%; Pred. No. 2.6e-25;
Matches 444; Conservative 250; Mismatches 735; Indels 572; Gaps 94;
QY 2 GAGSSTGQRPSPQAGSDTPSELVLSGHGPA-----AASGAADPPA-DADPATKLPQK 54
DB 550 GATSSSE-SEHQDQKSTEAFTSYDDIEPAKPTSSSEASGEGDVAKETTPAGE----- 604
QY 55 NGQLSSVGVAAE--OGDVHVOENQGEVEVDVQRESEDEVREKDRVEEMANSTA 112
DB 605 -----ASLAGEBEIVKGTTPAGEPSSSEG-DEETIVKGTTPAESSSSSE-----DELTKVTP 655
QY 113 VEDITKQDQGESEIIEIIPASENN---EEMVQ---PAESQANDVGKPKVFGFGFT 166
DB 656 AGEPSVAGEEELAK--ETTPAGEPSIAGEEELIVKTTPA-GSSSIAGEEEL-----VKVT 707
QY 167 VKDKNKSQDTYQLLTVKKDEGEAGASVAGADHQPSEVETVAGE--SASKSESLQSTE 224
DB 708 TPAGSSSSRGEBEIITKVTPPAGE-----SSSEGDDEIVVESTPAGPISEGEDVIAKITS 763
: : : : :

```


PA (UYOR-) UNIV OREGON HEALTH SCI.
 XX Klauck TM, Nauert JB, Scott JD;
 XX WPI, 1998-260552/23.
 DR
 XX
 PT New polypeptide fragments of protein kinase binding protein gravin -
 PT are useful for the study of modulation of action between gravin and
 PT protein kinase(s)
 XX
 PS Example 2; Column 43-46; 32pp; English.
 XX
 CC This sequence corresponds to residues 1130-1582 of gravin, and represents
 CC a polypeptide of the invention. The polypeptides are fragments capable of
 CC binding to type II regulatory subunit of cAMP-dependent protein kinase
 CC (PKA). Gravins is a kinase anchoring protein that binds to type II
 CC regulatory subunits of PKA and protein kinase C (PKC). Gravins is also an
 CC antigen of the autoimmune disease Myasthenia gravis (MG), where a patient
 CC develops antibodies against their own nicotinic acetylcholine receptors.
 CC The polypeptides are useful for providing analogues of gravin in the
 CC study of the modulation (e.g. blocking, inhibiting and stimulating) of
 CC interactions between gravin and kinase. The peptides are involved in the
 CC modulation of gravin-kinase interactions.
 CC
 XX Sequence 453 AA;
 SQ
 Query Match 7.9%; Score 639.5; DB 19; Length 453;
 Best Local Similarity 40.5%; Pred. No. 2.8e-26;
 Matches 183; Conservative 69; Mismatches 165; Indels 35; Gaps 16;
 OY 1110 VATCOV-----IKLOOL-MEQAAPSESETLTDSENGSTPLADSDPTADGTOODETDSQ 1163
 DB 6 VTTQCAETLAGVKSGEMVMEQAIIPDSVETPTDSETDSTVADFPDAGCTQKDEIYI 65
 OY 1164 DSKATAVROSQVTEEAATAQKEEPTLPNNVPAQEEHGEPPG-RDYLEPTQOELTAA 1222
 DB 66 EENEVHLV-PVNGTEAEAVPAQKERPPA-PSSFVFOEETKEQSKMEDTLEHTDKESVET 123
 OY 1223 VPVLAKTEVGOGEVDMLDGEKVK-----EEOEYFVHSG--PNSOKADVYTDSEVMGYA 1275
 DB 124 VSILSKTEGTQ--EADQYADEKTKDVPFEGLEGSGIDGTITVSREKVTVALKGEETEA 181
 OY 1276 GGOEKESTVOS-----LSLEGEMETDVEKEKETKPEQVSEEG-EOETAPBEHGTG 1329
 DB 182 ECKKDALELOSHAKSPSPVEREMVQVREKTEAEPTNVEEKLEHTAATVSEVSK 241
 OY 1330 KPVLTLDMPSSERKALGSLGGS-PSLPDQKAGCIEVVOISLDTVTQTAABAEKVI-- 1386
 DB 242 QLIQIVNPIIDGAKESVSLGSPPCLOGEBAVCTKIQVOSSEASFTLTAABEEKVIG 301
 OY 1387 ETVVISSEGESPECVGAHLLEAKSSATGCHMTLOHAEPTVPLGEGSAESIPITVPAP 1446
 DB 302 ETANILETGETLEPAGAHVLEBKSSSEKNEDEFAHNGEDAVPTGPDCAKSTPVIVSAT 361
 OY 1447 ESTLHPDLOGEISASQRESEEDKPDADGKSTAIKVLKABE--ILLEESKSN 1504
 DB 362 KKGSLSDLEGETTSLKMKSDVDEQVACQEV--KVSVAIEDL--EPENGILELETKSS 416
 OY 1505 KIVLVNIOTAVDOFART-ETAPETHAYDSQTQ 1535
 DB 417 KLVONIIOTAVDOFVTEETATEMLTSELQTQ 448

RESULT 14
 AAB18407
 ID AAB18407 standard; protein; 453 AA.
 XX
 AC AAB18407;
 XX
 DT 26-JAN-2001 (first entry)
 XX
 DE Human gravin PKA regulatory subunit RII binding site assay peptide #1.

KM Human; gravin; PKA RII binding site; myasthenia gravis;
 KM kinase anchoring protein; cAMP dependent protein kinase.
 XX
 OS Homo sapiens.
 XX
 PN US6090929-A.
 XX
 PD 18-JUL-2000.
 XX
 PF 19-DEC-1997; 97US-0994570.
 XX
 PR 19-DEC-1996; 96US-0769309.
 XX
 PA (UYOR-) UNIV OREGON HEALTH SCI.
 XX Klauck TM, Scott JD, Nauert JB,
 XX WPI, 2000-523763/47.
 DR
 XX
 PT Novel polynucleotides useful for detecting gravin in patients suffering
 PT from Myasthenia gravis encodes cAMP-dependent protein kinase-binding
 PT polypeptide and protein kinase C-binding polypeptide of gravin -
 CC
 PS Example 2; Column 47-50; 34pp; English.
 XX
 CC The present sequence is part of the human gravin protein. Gravins is an
 CC A-kinase anchoring protein (AKAP) which is involved in the localisation
 CC of cAMP dependent protein kinase A (PKA) via interactions between the
 CC RII binding region and the PKA regulatory subunit RII. Gravins is also an
 CC antigen found in myasthenia gravis sufferers, and it is thought that
 CC antibodies to it may be useful in modulating the binding of PKA, and thus
 CC aid in the treatment of the disease. This sequence was used in an assay
 CC to identify the precise position of the RII binding site within the
 CC gravin protein.
 CC
 XX Sequence 453 AA;
 SQ
 Query Match 7.9%; Score 639.5; DB 21; Length 453;
 Best Local Similarity 40.5%; Pred. No. 2.8e-26;
 Matches 183; Conservative 69; Mismatches 165; Indels 35; Gaps 16;
 OY 1110 VATCOV-----IKLOOL-MEQAAPSESETLTDSENGSTPLADSDPTADGTOODETDSQ 1163
 DB 6 VTTQCAETLAGVKSGEMVMEQAIIPDSVETPTDSETDSTVADFPDAGCTQKDEIYI 65
 OY 1164 DSKATAVROSQVTEEAATAQKEEPTLPNNVPAQEEHGEPPG-RDYLEPTQOELTAA 1222
 DB 66 EENEVHLV-PVNGTEAEAVPAQKERPPA-PSSFVFOEETKEQSKMEDTLEHTDKESVET 123
 OY 1223 VPVLAKTEVGOGEVDMLDGEKVK-----EEOEYFVHSG--PNSOKADVYTDSEVMGYA 1275
 DB 124 VSILSKTEGTQ--EADQYADEKTKDVPFEGLEGSGIDGTITVSREKVTVALKGEETEA 181
 OY 1276 GGOEKESTVOS-----LSLEGEMETDVEKEKETKPEQVSEEG-EOETAPBEHGTG 1329
 DB 182 ECKKDALELOSHAKSPSPVEREMVQVREKTEAEPTNVEEKLEHTAATVSEVSK 241
 OY 1330 KPVLTLDMPSSERKALGSLGGS-PSLPDQKAGCIEVVOISLDTVTQTAABAEKVI-- 1386
 DB 242 QLIQIVNPIIDGAKESVSLGSPPCLOGEBAVCTKIQVOSSEASFTLTAABEEKVIG 301
 OY 1387 ETVVISSEGESPECVGAHLLEAKSSATGCHMTLOHAEPTVPLGEGSAESIPITVPAP 1446
 DB 302 ETANILETGETLEPAGAHVLEBKSSSEKNEDEFAHNGEDAVPTGPDCAKSTPVIVSAT 361
 OY 1447 ESTLHPDLOGEISASQRESEEDKPDADGKSTAIKVLKABE--ILLEESKSN 1504
 DB 362 KKGSLSDLEGETTSLKMKSDVDEQVACQEV--KVSVAIEDL--EPENGILELETKSS 416
 OY 1505 KIVLVNIOTAVDOFART-ETAPETHAYDSQTQ 1535
 DB 417 KLVONIIOTAVDOFVTEETATEMLTSELQTQ 448


```

Db 4201 EEBQOLAIVSESEPAVNEVEVSEAPF-----SKIYE-----BEVIAIEKPKKEFT 4245
QY 1136 ----DSETNSTPLADSDTADGTQODET-----IDSQSKATAVRO----- 1173
Db 4246 IRVSESEBKPPEEPSEVQFTVKRRKPSVTFADEBATEIIVIKESKPAEAVVTEDAHIKTKKPK 4305
QY 1174 SQVTEBEAATAQKEPSTLPNNVPAQEHGSEBGRDVLPTQOELTAANVPYLAKT-EVG 1232
Db 4306 KKVTDVEAEELKIKITEEVPQEIPLIEVSEB--EVITETIK--TAPVEEKTYSKIG 4358
QY 1233 QEGEVDWMDGKEVKEQEVFVHSGPNSQKADAVTYDSEVMGVAGCQEKESTEVQSLSEB 1292
Db 4359 IKETPEKPAIAIYEEEPVVTPEIEBAPKEVFEHKKVRI-----EETPRELVE 4409
QY 1293 GEMETDYKEKERTKPEQVSEGEQETAPEHEGTYGKPVLTLMPSSEBKGALGSLGGS 1352
Db 4410 EVIEEEVAVIRKKKPKPEIKEEPEAEVTV-----STPKPV-----EVEEATSSIAVI 4456
QY 1353 PSLPDQDAGCIEVVOVQSLDTVTQTAFAVEKIVETVISE-----TGESP 1398
Db 4457 PEQPTEEEAADLKITIEE--TPQELVOEIEIEIEIEVEKPAEEOPTDFTFATKQSE 4513
QY 1399 ECVGAHLPAEKSSATGHWTLQHAEDTVPLG--PESQAESIPITVTPAPESTLHPDLQ 1455
Db 4514 KKPVTVEELPEEVO-----TIQKKKKAPVEVEEPEAEFLVVKPTVQEVTV-----BE 4562
QY 1456 GEISASQRESESEEDKPDAGPDADGKESTAIEKVLKABEILEL-----ESKSNKIY 1507
Db 4563 AKITSKKPKVKEE-----AAAEKVTITEEIPTEBEVOEIIIEIIEIEEKPAEYV 4614
QY 1508 LNVIOATAVDQPARTET-----APETHAVDSQ-TQVPACRLDSREPRRCWTMKDAKM 1558
Db 4615 IEVKESQPEAVEDEKVSLLPKKKPKAPIVEEPEAEITLKKVKSSEVOEBAKIVKKKPKKI 4674
QY 1559 -----KHPVQOPREDLOVLTVEAWAOPKRCULPR-----LOLKAP 1593
Db 4675 DEVAVADELTVKVEEVPPEPIVEEVIIEEPIKPKPEPEPEDIVDAAIVLKLKP 4730

```

Search completed: December 13, 2002, 00:12:16
 Job time : 141 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 13, 2002, 00:10:50 ; Search time 38 Seconds
(without alignments)
1235.763 Million cell updates/sec

Title: US-09-902-432-4
Perfect score: 8073
Sequence: 1 MGAGSTEQSPQAGSDT.....AWAQRKCLRLQLKAPVSK 1596

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5421.5	67.2	1346	2	US-08-635-121-2
2	4274.5	52.9	1780	1	US-08-769-309A-5
3	4274.5	52.9	1780	3	US-08-994-570-5
4	952.5	11.8	292	1	US-08-769-309A-3
5	952.5	11.8	292	3	US-08-994-570-3
6	656	8.1	651	1	US-08-769-309A-17
7	656	8.1	651	3	US-08-994-570-17
8	639.5	7.9	453	1	US-08-769-309A-14
9	639.5	7.9	453	3	US-08-994-570-14
10	521.5	6.5	396	1	US-08-769-309A-15
11	521.5	6.5	396	3	US-08-994-570-15
12	462	5.7	1786	4	US-08-973-462-8
13	388	4.8	1805	1	US-07-853-913-2
14	365.5	4.5	3696	4	US-09-134-001C-5080
15	361.5	4.5	2101	1	US-08-466-390-4
16	361.5	4.5	2101	1	US-08-470-950-4
17	361.5	4.5	2101	1	US-08-467-781-4
18	361.5	4.5	2101	2	US-08-483-924-4
19	361.5	4.5	2101	4	US-09-452-294-1
20	361	4.5	2137	4	US-09-134-001C-4463
21	358.5	4.4	2101	1	US-08-195-487-4
22	358.5	4.4	2101	5	PCT-US93-06160-4
23	354	4.4	1618	1	US-07-853-913-4
24	354	4.4	2409	6	5180808-2
25	334.5	4.1	1162	2	US-08-728-323A-2
26	334.5	4.1	1162	4	US-09-238-568-2
27	334	4.1	1018	1	US-08-072-610-2

28 334 4.1 1018 2 US-08-719-822B-2 Sequence 2, Appli
29 334 4.1 1018 4 US-09-092-458-2 Sequence 2, Appli
30 321 4.0 1852 1 US-08-425-061-24 Sequence 24, Appli
31 321 4.0 1852 2 US-08-825-886-24 Sequence 24, Appli
32 321 4.0 1863 1 US-08-425-061-16 Sequence 16, Appli
33 321 4.0 1863 1 US-08-480-784-2 Sequence 2, Appli
34 321 4.0 1863 1 US-08-483-553-2 Sequence 2, Appli
35 321 4.0 1863 1 US-08-487-002-2 Sequence 2, Appli
36 321 4.0 1863 1 US-08-483-554B-2 Sequence 2, Appli
37 321 4.0 1863 1 US-08-488-011B-2 Sequence 2, Appli
38 321 4.0 1863 2 US-08-825-886-16 Sequence 16, Appli
39 321 4.0 1863 4 US-08-850-727-2 Sequence 2, Appli
40 321 4.0 1863 5 PCT-US95-10202-2 Sequence 2, Appli
41 321 4.0 1863 5 PCT-US95-10203-2 Sequence 2, Appli
42 321 4.0 1863 5 PCT-US95-10200-2 Sequence 2, Appli
43 320 4.0 1898 1 US-08-056-200-94 Sequence 94, Appli
44 320 4.0 1898 2 US-08-800-644-94 Sequence 94, Appli
45 319.5 4.0 1939 4 US-09-310-187A-1 Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-635-121-2
; Sequence 2, Application US/08635121
; Patent No. 5910442
; GENERAL INFORMATION:
; APPLICANT: Gelman, Irwin H.
; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10112-0228
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/635,121
; FILING DATE: 19-APRIL-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Richard S
; REGISTRATION NUMBER: 26,154
; REFERENCE/DOCKET NUMBER: A30558 - 165/33603
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-408-2558
; TELEFAX: 212-765-2519
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1346 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
US-08-635-121-2

Query Match 67.2%; Score 5421.5; DB 2; Length 1346;
Best Local Similarity 91.6%; Pred. No. 0;

Matches 1098; Conservative 24; Mismatches 62; Indels 15; Gaps 7;			
Qy	387	MEAHQEVAAVAVSTVEKTEEEGGGGGAEAGVVEGTGSELPEPKLAEPQEVQEAEP	446
Db	1	MEAHQEVAAVAVSTVEKTEEEGGGGGAEAGVVEGTGSELPEPKLAEPQEVQEAEP	60
Qy	447	ELMKSRMCMVSGGDHQTLPDSPEKTLPHKHPREGIVSEVMLMSQERIKVQSGSKLFL	506
Db	61	ELMKSRMCMVSGGDHQTLPDSPEKTLPHKHPREGIVSEVMLMSQERIKVQSGSKLFL	120
Qy	507	SSGLKTLGKKGKKGKGGGDEEPEGYOHITESPESADQKSGSSASPEPEPTTCL	566
Db	121	SSGLKTLGKKGKKGKGGGDEEPEGYOHITESPESADQKSGSSASPEPEPTTCL	180
Qy	567	EKGPLAPQDEAEAGTTSDEKKRE---GITPMASFKNMTPKRVRPSESDKEEL	622
Db	181	EKGPLAPQDEAEAGTTSDEKKRE---GITPMASFKNMTPKRVRPSESDKEEL	234
Qy	623	EKVSATLSTDSVSEMDKTVGEQKPEEPKRRVDTSVSMALLCVGSSKKRARKA	682
Db	235	EKVSATLSTDSVSEMDKTVGEQKPEEPKRRVDTSVSMALLCVGSSKKRARKA	294
Qy	683	SSSDDEGPRTLGGDSHAEAKSKDEAGTAVPASTQEQDQAGSSPAPSGSEGG	742
Db	295	SSSDIR-GPRTLGGGQSGSRGGQQRRTDAVPASTQEQDQAGSSPAPSGSEGG	353
Qy	743	VSTWSEFKRLVTPRKSKSKLEEK-AEDSVYQSLTEIPEPSRESWVS-IKKPIGRK	799
Db	354	VSTWSEFKRLVTPRKSKSKLEEKAGRTLVYAGCPLRSNRVEKNLGFPLRNSPDGGR	413
Qy	800	KRAGQKQATVDESGPFIENEDDNNPVPVPLSTYNAVEEKMAQNTLPLQAGVY	859
Db	414	KQMGROQATVDESGPFIENEDDNNPVPVPLSTYNAVEEKMAQNTLPLQAGVY	472
Qy	860	VSEELSKTLVTVSAVVDGTRAVTVSEERSPSWISASVTEPLETAGEAMPVEVEEK	919
Db	473	VSEELSKTLVTVSAVVDGTRAVTVSEERSPSWISASVTEPLETAGEAMPVEVEEK	532
Qy	920	DIIEETPVLTQTLPEGADHDVMTSEVDTSEAVTATETSEALRTEVTEASAEETT	979
Db	533	DIIEETPVLTQTLPEGADHDVMTSEVDTSEAVTATETSEALRTEVTEASAEETT	592
Qy	980	DMVAVSOLTPSPDTTEATPVQEVESGVLDTSEERQTOALQVADKVEESQVPAQ	1039
Db	593	DMVAVSOLTPSPDTTEATPVQEVESGVLDTSEERQTOALQVADKVEESQVPAQ	652
Qy	1040	TVORTGSALEKVEVEEDSEVLASEKEKDVMPKGPVOEAGAEHLAOGSETQATPESLE	1099
Db	653	TVORTGSALEKVEVEEDSEVLASEKEKDVMPKGPVOEAGAEHLAOGSETQATPESLE	712
Qy	1100	VPEVTADVHVATCOVILQOLMEQAVAPESSETLTDSETNGSTPLADSDTADGTQDET	1159
Db	713	VPEVTADVHVATCOVILQOLMEQAVAPESSETLTDSETNGSTPLADSDTADGTQDET	772
Qy	1160	IDSQDSKATAVROSQVTEEEBAATQKEEPSTLPNNVPAQEHGEBPRDVLLEPTQDELA	1219
Db	773	IDSQDSKATAVROSQVTEEEBAATQKEEPSTLPNNVPAQEHGEBPRDVLLEPTQDELA	832
Qy	1220	AAAPVPLAKTEVGOGEVDWLDGEKVEKEOEYFVHSGNSQKADVTADSEWVGAQOE	1279
Db	833	AAAPVPLAKTEVGOGEVDWLDGEKVEKEOEYFVHSGNSQKADVTADSEWVGAQOE	892
Qy	1280	KESTEVOGLSIEGEMETDVEKEKETKPEQVSEGEQETAPEHEGTYGKVLTLDMPS	1339
Db	893	KESTEVOGLSIEGEMETDVEKEKETKPEQVSEGEQETAPEHEGTYGKVLTLDMPS	952
Qy	1340	SERKALGSLGSGPSLPQDRACTIEVOVQSLDTTDTVTAENAVETVITVISTGESP	1399
Db	953	SERKALGSLGSGPSLPQDRACTIEVOVQSLDTTDTVTAENAVETVITVISTGESP	1012
Qy	1400	CVGAHLPLAEKSSATGMMTLOHAEDTVPLGPEQAESPIITVTPAPBSTLHPDLQGFIS	1459
Db	1013	CVGAHLPLAEKSSATGMMTLOHAEDTVPLGPEQAESPIITVTPAPBSTLHPDLQGFIS	1072

Qy	1460	ASQRESEEDKPDAGDADGKESATKYLKAPETLLELSKSNKTVLVITQAVDQFA	1519
Db	1073	ASQRESEEDKPDAGDADGKESATKYLKAPETLLELSKSNKTVLVITQAVDQFA	1132
Qy	1520	RTEPAPETHAVDSQVPAICRLDSREPNRCMTKMKDKMKHPVQPREDDQVITLBEAM	1578
Db	1133	RTEPAPETHAVDSQVPAICRLDSREPNRCMTKMKDKMKHPVQPREDDQVITLBEAM	1191
RESULT 2			
US-08-769-309A-5			
Sequence 5, Application US/08769309A			
Patent No. 5741890			
GENERAL INFORMATION:			
APPLICANT: Scott, John D.,			
APPLICANT: Nauer, Brian J.,			
APPLICANT: Klaus, Theresa M.			
TITLE OF INVENTION: Protein Binding Domains of Gravin			
NUMBER OF SEQUENCES: 24			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun			
STREET: 6300 Sears Tower/233 South Wacker Drive			
CITY: Chicago			
STATE: Illinois			
COUNTRY: United States of America			
ZIP: 60606-6402			
COMPUTER READABLE FORM:			
MEDIUM TYPE: Floppy disk			
COMPUTER: IBM PC compatible			
OPERATING SYSTEM: PC-DOS/MS-DOS			
SOFTWARE: Patentin Release #1.0, Version #1.30			
CURRENT APPLICATION DATA:			
APPLICATION NUMBER: US/08/769,309A			
FILING DATE:			
CLASSIFICATION: 435			
ATTORNEY/AGENT INFORMATION:			
NAME: No. 5741890and, Greca E.			
REGISTRATION NUMBER: 35,302			
REFERENCE/DOCKET NUMBER: 27866/33451			
TELECOMMUNICATION INFORMATION:			
TELEPHONE: 312-474-6300			
TELEFAX: 312-474-0448			
TELEX: 25-3856			
INFORMATION FOR SEQ ID NO: 5:			
SEQUENCE CHARACTERISTICS:			
LENGTH: 1780 amino acids			
TYPE: amino acid			
TOPOLOGY: linear			
MOLECULE TYPE: protein			
US-08-769-309A-5			
Query Match 52.9%; Score 4274.5; DB 1; Length 1780;			
Best Local Similarity 57.8%; Pred. No. 9.2e-255;			
Matches 957; Conservative 187; Mismatches 402; Indels 109; Gaps 36;			
Qy	1	MGASSSTGQSRPEQ-PAGSDTSELVLSGHQPAEAS-GAAGDPADA--DPATKLPQKNG	56
Db	1	MGASSSTGQSRPEQ-PAGSDTSELVLSGHQPAEAS-GAAGDPADA--DPATKLPQKNG	60
Qy	57	QLSSVNGVABGDVHVGEBNDEGQ-----EEEVNDEVDYQGRSEVREKDR	102
Db	61	QLSSVNGVABGDVHVGEBNDEGQ-----EEEVNDEVDYQGRSEVREKDR	120
Qy	103	VEEMAANSTAVEDITTKQGOEETSEITIOIPASENNVEMVQPAESQANDVGFKKVPKPVG	162
Db	121	VEEMAANSTAVEDITTKQGOEETSEITIOIPASENNVEMVQPAESQANDVGFKKVPKPVG	179
Qy	163	FKFTVKDKNEKSDTVQLTLTKKDEGEGAEASVAGDHOEPPSVETAVGESAKSESKLOS	222
Db	180	FKFTVKDKNEKSDTVQLTLTKKDEGEGAEASVAGDHOEPPSVETAVGESAKSESKLOS	234
Qy	223	TEKQEGTLKQEQSTELPLQAESQDAABEAKDEGEEKQEKPTKSPSPSPVNSSTTS	282

Db 235 TEKPEETLKREQSHAEISPPAESQAV--BECKEGERKEQKPKSAESPTSPVTSGETS 293
Qy 283 SFKKPFTHWAGWKKTSFKKSKEDDLLETABKRXEQEAKVDEBEKTEPASEE----- 337
Db 294 TFKKFFFTQWAGWKKTSFRKPKEDVEASESKKXEQEPKVDTEEDGKAEVASEKLTASE 353
Qy 338 ---OEPAEDTDQARLSADYKVLPLPDQVDLEASSEKCAPLATEVDFEKMEHQ-E 392
Db 354 QAHPOEPAESAHPRLSAEYKVLPEBEEQVSGSQGPSEKPPAPLATEVDFEKEVHQEE 413
Qy 393 VVAEYHVSTVEKTEBEEQGGGAEAGVVEGTGSLPEKLAEPQVPOEPAEAEELMKS 452
Db 414 VVAEYHVSTVEETEQ-----KTEVEETAGSVPAEELVGMDAEPQOEAEPKELVKL 465
Qy 453 REMCVSGDHTQLDLSPKEETKPKHPEGIVSEVEMLSQBRIRIKVQSPKPKLFSSGLK 512
Db 466 KETCVSGEDPTQADLSPDEKVLSPKPEGVVSEVEMLSQBRIRIKVQSPKPKLFSSGLK 525
Qy 513 KLSGKKOKRGKGGGDEBPGEYQIHHTESPESADEQKGESSASSPPEETTCLEKPLE 572
Db 526 KLSGKKOKGRK--GGDESGEHTQVPADSPDSQEQKGESSASSPPEETTCLEKGLAE 584
Qy 573 APODGEAEETTSQGEKKREGITPWASKMVTPKKVRPSESDKGEELKVKSATLSS 632
Db 585 VQDGEAEAGTSQGEKKREGITPWASKMVTPKKVRPSESDKGEELKVKSATLSS 644
Qy 633 TDSTVSEMDQEVKTVGBEQKPEERKRVDTSVSWEALICVGSKKRKARKASSDDEGGPR 692
Db 645 TESTASEMEQEMKGSVEEPKPEERKRVDTSVSWEALICVGSKKRRARRSSDEEGPK 704
Qy 693 TLGDSHRAEASXDKAGTDVAPASTQEQOQAGSSPEPAGSPSGEGVSTWESPKRL 752
Db 705 AMGDDHOKADGAKDKETGTDIGLAGQEHDPGQGSSEPEQAGSPTEGEGVSTWESPKRL 764
Qy 753 VTPRKSKSLEEKAEQ-----SSVEQLSTETEPSEBSWSIKFICRRKKRAGDKQEQ 808
Db 765 VTPRKSKSLEEKSEOSIAGSGVEHSTPDTEPGKESWSIKFIPGRRKKRPDQKQEQ 824
Qy 809 ATVEDSGPVEINEDDPNPAVPLSEYNAVERKME---AQNTPELPOLLGAVVSVBELS 865
Db 825 APVEDAGPTGANEDSDVPAVPLSEYDAVERKMEQAQKGAQEQKAAETVSKELS 884
Qy 866 KTLVHTVSVAVIDGTRAVTSVEERSPWISASVTEPLEHTAGEAMPVVEVTEKDIIA-E 924
Db 885 ESQVHMAAAVADGTRAATIIERSPSWISASVTEPLEQVEAEAAALLTEEVLEEVIAEE 944
Qy 925 ETPVLTOTLPEGKADHDMVTSEVDFTSEAVTATETSEALRTEETVTEASGAEETDMVSA 984
Db 945 EPPTVTEPLPENREARGDVTVSEALTEPAVTAATAETAGPLGSEEGTEASAAETTEMVSA 1004
Qy 985 VSQLTSDPTTEATPVQESVGLDTEEBEROTQAILQAVADKVKESQVPAE----- 1038
Db 1005 VSQLTSDPTTEATPVQESVGLDTEEBEROTQAILQAVADKVKESQVPAE----- 1064
Qy 1039 -QTVORTGSKALEKVEEVEDSEVLASEKEKXDMKPGVQVQAGAEHLAQGSSETQCATPES 1097
Db 1065 LQPVQRA-----EAERPEQAEASGLKCTETDVLVKVDAQEAKTEPFTQGVVGTTPES 1118
Qy 1098 LE-VPEVTADVH---VATQV-----IKLOOL-MEQAVAPESSETTDSNKGSTPLAD 1147
Db 1119 FEKAPQVTESESSELVTTQCAETLAGVKSQEMVMEQAIIPDSVETPTDSETDGSSTPAD 1178
Qy 1148 SDTADGTQODTIBSQSKATAAARQSVQVTEEAATAQKEPSTLPNNVPAQEBEGREP 1207
Db 1179 FDAGTTQKDBIVIHRENEVHLV--PVRGTBAEAVPAQKERRPPA-PSFVFQEBTKQSK 1236
Qy 1208 -RDVLEPTQOELTAAAVPVLAKTEVQGEVGDWLDGKVK-----ERQEVFVHSG--PNS 1259
Db 1237 MEDTLEHTDKREVSVETVILSKTEGTQ--EADQYADEKTKDVPFEGLEGSIDGTIVSR 1294
Qy 1260 QKAADVTYDSEVMGACQEKESTEVOS-----LSLEGEEMETDVEKEKRETKPEQVSEE 1314

Db 1295 EKYTEVALKGEETEEAECKDDALELOSHAKSPSPSPVEREMVQVBEREKTEAPETHNEE 1354
Qy 1315 G-BOETAAPHEGTYGKPVLTLDMPSSERKALGSLGGS--PSLPDQDKAGCIEVQVQSLD 1372
Db 1355 KLEHETAVTVSEVSKQLQTVNVPIDGAKVSSLSGSPPPCLGQBEAVCTKIQVQSSSE 1414
Qy 1373 TTVTQTAEAVEKVI--ETVVISSETGESPECVGAHLLPAEKSSATGGHWTLOHAEDTVPLG 1430
Db 1415 ASFTLTAAABEEKVLGTETANILETGETLEPAGAHVLEEKSSSEKNEFAAHPGEDAVPTG 1474
Qy 1431 PESOAESPIIIVTPAPESTLHPDLOGEISASQERSERSEEDKPDAGPDADGKESFAIEKVL 1490
Db 1475 PDQAKSTPVIVSATYKGLSSDLGEKTTSLKWSDEVDEQVACQEV--KVSVAIEDL- 1531
Qy 1491 KABPE--ILEBSKSNKVLNVIQTAVDQPART-ETAPETHAYDSOTQVPCRLDSREPN 1547
Db 1532 --BPENGILSELTSSKLVQNIQTAVDQVFRTEETATEMLTSELQTOAHVIRKADSDAG 1589
Qy 1548 RCWTK-----MKDAKMKHPVPQPREDLQVLTVLEA 1577
Db 1590 QETEKEGEPQASQADETPITSKESESESTAVGQA 1624

RESULT 3
US-08-994-570-5
; Sequence 5, Application US/08994570
; Patent No. 6090929
; GENERAL INFORMATION:
; APPLICANT: Scott, John D.,
; APPLICANT: Nauert, Brian J.,
; APPLICANT: Klauack, Theresa M.,
; TITLE OF INVENTION: Protein Binding Domains of Gravin
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower/233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/994,570
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6090929and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/33451
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1780 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-994-570-5

Query Match 52.9%; Score 4274.5; DB 3; Length 1780;
Best Local Similarity 57.8%; Pred. No. 9.2e-255;
Matches 957; Conservative 187; Mismatches 402; Indels 109; Gaps 36;
Qy 1 MGAGSSTEQRSPEQ-PAGSDTPSELVLSGHGPAEAS-GAAGDPADA--DPATKLPKNG 56
Db 1 MGAGSSTEQRSPEQPPGSGSTPAEPSPGSGPAAEAPDTTADPAIAASDPATKLPKNG 60

QY 57 QLSVNGVABEGDVHVOENBQ-----EEVVDDEVGORSEEDREKOR 102
 Db 61 QLSVNGVABEGDVHVOENBQ-----EEVVDDEVGORSEEDREKOR 120
 QY 103 VEEAANSTAVEDTJKOQOETSEIIEOI PASENNVEEMVOPAESQADVGFKKVFKVVG 162
 Db 121 DKEAATKSAVVDITDDQOENRN-IEOI PSESNSLLELTQPTESQANDIGFKKVFVVG 179
 QY 163 FKFTVKDKNEKSDTVOLLTVKKDGEAGASVAGDHQEPSVETAVGASAKESSELKOS 222
 Db 180 FKFTVKDKNEKSDTVOLLTVKKDGEAGASVAGDHQEPSVETAVGASAKESSELKOS 234
 QY 223 TEKQOBTJKOQOETSEIIEOI PASENNVEEMVOPAESQADVGFKKVFKVVG 282
 Db 235 TEKQOBTJKOQOETSEIIEOI PASENNVEEMVOPAESQADVGFKKVFKVVG 293
 QY 283 SPKFFFTGMGMRKKTFFKSKEDDLTAEARKQOAKVDEEKEKTEPASEE----- 337
 Db 294 TPKKFTGMGMRKKTFFKSKEDDLTAEARKQOAKVDEEKEKTEPASEE----- 353
 QY 338 ----OEPAEDTDQARLSADYKVELPLBEDQVGLLEASSEKCAPLATEVDEKMAHQ-E 392
 Db 354 QAHPEPASAHEPRLSAEYKVELPSEBQVGSQGPSEKPAPLATEVDEKMAHQ-E 413
 QY 393 VVAEYHVSVEKTEBEGGSGGAEAGVYEGESLPEPKLAEPQVQOAPAEELMKKS 452
 Db 414 VVAEYHVSVEKTEBEGGSGGAEAGVYEGESLPEPKLAEPQVQOAPAEELMKKS 465
 QY 453 REMCVSGDHTQULDTLSPEKTLPHNPEGIVSEVEMLSQERIKYQGSPLKLFSSGLK 512
 Db 466 KETCVSGDHTQULDTLSPEKTLPHNPEGIVSEVEMLSQERIKYQGSPLKLFSSGLK 525
 QY 513 KLSGKKQKRGCGGDEBEGEYQIHTESPESADEQKSGSSASSPPEPTECCKEPL 572
 Db 526 KLSGKKQKRGCGGDEBEGEYQIHTESPESADEQKSGSSASSPPEPTECCKEPL 584
 QY 573 APQDEABEGTSDDEKREGITPMASPKKMTPKKRRPSESPKEBELVKSATLSS 632
 Db 585 APQDEABEGTSDDEKREGITPMASPKKMTPKKRRPSESPKEBELVKSATLSS 644
 QY 633 TDSVTSEMODEKTVGEOBKPEPKRRVDTSVSWEALICVSGSKRRARASSDDEGPR 692
 Db 645 TDSVTSEMODEKTVGEOBKPEPKRRVDTSVSWEALICVSGSKRRARASSDDEGPR 704
 QY 693 TLGDSHRAEASKDEAGTDAVPASTOQOAGSSSPBAGSPBEGGVSTWESFKKL 752
 Db 705 TLGDSHRAEASKDEAGTDAVPASTOQOAGSSSPBAGSPBEGGVSTWESFKKL 764
 QY 753 VTPRKKSXSLEKKAED----SSVEQLSTEIPRSRESVNSIKKFLPGRKKRAGQKQO 808
 Db 765 VTPRKKSXSLEKKAED----SSVEQLSTEIPRSRESVNSIKKFLPGRKKRAGQKQO 824
 QY 809 ATVEGSPETINEDDPNVAVVPLSEFYNAVEBEKME--ACQNTLPLGLAVVYSELS 865
 Db 825 ATVEGSPETINEDDPNVAVVPLSEFYNAVEBEKME--ACQNTLPLGLAVVYSELS 884
 QY 866 KTLVHVSVAVIDGTRAVTVSEERSPSWISASVTEPLEHTAGEAMPVEVTEKDIIA-E 924
 Db 885 KTLVHVSVAVIDGTRAVTVSEERSPSWISASVTEPLEHTAGEAMPVEVTEKDIIA-E 944
 QY 925 ETPVLTQTLPEGKDAHDMVTSVDFTSAAVATETSEALRTBEVYEAAGAETDMVA 984
 Db 945 ETPVLTQTLPEGKDAHDMVTSVDFTSAAVATETSEALRTBEVYEAAGAETDMVA 1004
 QY 985 VSQULDSPTTEAPRVPVQESGVLDTEEBERQTOAILAVVADKVEEEOVAT----- 1038
 Db 1005 VSQULDSPTTEAPRVPVQESGVLDTEEBERQTOAILAVVADKVEEEOVAT----- 1064
 QY 1039 -QTVORTGSKALEKVEEVEDESVLASEKXKDVMPKQVQOAGAEHLAQSSETQATPES 1097
 Db 1065 -QTVORTGSKALEKVEEVEDESVLASEKXKDVMPKQVQOAGAEHLAQSSETQATPES 1118
 QY 1098 LE-VPEVTADVDH---VATQV-----IKLQOL-MEQAVAPESSELTLDSENGSTPLAD 1147

Db 1119 PEKAPQVTESTESSELVTTQOAEETLAVGKQOEMMEQALIPDVSJETDTSDSTVAD 1178
 QY 1148 SPTADGTQOEDTSDSOKATPAVROQVTEEBAAVQKEEPSTLPNNVPAQOEHEGEPG 1207
 Db 1179 SPTADGTQOEDTSDSOKATPAVROQVTEEBAAVQKEEPSTLPNNVPAQOEHEGEPG 1236
 QY 1208 -RDVLEPQOELTAAAPVLAKTVEQGEVDMLDGKVK-----EQEYFVHSG--PNS 1259
 Db 1237 MEDTLEHTDKESVSEVTSILSKTEGTQ--EADQYADEKTDVPPFEGLEGSIDGITVSR 1294
 QY 1260 OKADVTYDSEVWAGVQOKEESTEVOS-----LSLEGEEMETVEKEKTEKPEOYSEE 1314
 Db 1295 EKVTEVALKSGTEBAECKODDALELOSHAKSPSPPEREMVAVVEKEKTEBAETHNEE 1354
 QY 1315 G-EQETAPHEGTYGKPVLTLDMPSSERKALGSLGS--PSLPDQKACIEVOVOSLD 1372
 Db 1355 KLEHETAVTVEEVSQOLLQVNVPIIDGAEVSSLSGSPPCGQEEAVCTKIQVQOSE 1414
 QY 1373 TTVVQTAFAVKVI--ETVVISSEGESPECVGAHLPAEKSSATGHWITLOHADTVPLG 1430
 Db 1415 ASFTVTAABEKEKVLGTANTLETGETLEPAGAHVLEEKSSKXNEDFAHPGDDAVPTG 1474
 QY 1431 PESQAESIPITVPAPASTLHPDLQGEISASQERSEEDKPDAGPDAGKESTAIKVL 1490
 Db 1475 PDQAKSTPVIVSATTKKGLSSDLEGEKTSILKXKSDEVDQVACQEV--KVSVAIEDL- 1531
 QY 1491 KAEBE--LLEESKSNKIVLNVIOAVDOFART-ETAPETHAYDSQVOPACRLDSREP 1547
 Db 1532 --EPENGILEETKSSKLQVNIIOAVDOFARTETETEMLTSELQOAHVIKADSSQAG 1589
 QY 1548 RCWTK-----MKDAMKHPVPQREDLOVLTVLEA 1577
 Db 1590 QETKEGEEPOASQADETPITSAKESSESTAVGQA 1624

RESULT 4
 US-08-769-309A-3
 / Sequence 3, Application US/08769309A
 / Patent No. 5741890
 / GENERAL INFORMATION:
 / APPLICANT: Scott, John D.,
 / APPLICANT: Nauert, Brian J.,
 / APPLICANT: Klauack, Theresa M.,
 / TITLE OF INVENTION: Protein Binding Domains of Gravin
 / NUMBER OF SEQUENCES: 24
 / CORRESPONDENCE ADDRESS:
 / ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
 / STREET: 6300 Sears Tower/233 South Wacker Drive
 / CITY: Chicago
 / STATE: Illinois
 / COUNTRY: United States of America
 / ZIP: 60606-6402
 / COMPUTER READABLE FORM:
 / MEDIUM TYPE: Floppy disk
 / COMPUTER: IBM PC compatible
 / OPERATING SYSTEM: PC-DOS/MS-DOS
 / SOFTWARE: Patentin Release #1.0, Version #1.30
 / CURRENT APPLICATION DATA:
 / APPLICATION NUMBER: US/08/769,309A
 / FILING DATE:
 / CLASSIFICATION: 435
 / ATTORNEY/AGENT INFORMATION:
 / NAME: No. 5741890and, Greta E.
 / REGISTRATION NUMBER: 35,302
 / REFERENCE/DOCKET NUMBER: 27866/33451
 / TELECOMMUNICATION INFORMATION:
 / TELEPHONE: 312-474-6300
 / TELEFAX: 312-474-0448
 / TELEX: 25-3856
 / INFORMATION FOR SEQ ID NO: 3:
 / SEQUENCE CHARACTERISTICS:
 / LENGTH: 292 amino acids

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-769-309A-3

Query Match 11.8%; Score 952.5; DB 1; Length 292;
Best Local Similarity 66.1%; Pred. No. 2.5e-51;
Matches 199; Conservative 30; Mismatches 53; Indels 19; Gaps 4;

QY 254 KDGEEKQKEPTKSPSPVNSSETTSFKFTHGWAGWRKKTSPKSKEDDLTAE 313
DB 1 KEEGEEKQKEPSKSAESPTSPVTSSTGTFKFFTCQWAGWRKKTSPKSKEDDEVEASE 60
QY 314 KRKEQAEKVDEEKEKTEPASEE-----QEPADTDQARLSADYKVELPLEDOV 364
DB 61 KKEQEPKVDTEEDGKAESAELKATASEQAHPQEPASAEHPRLSAEYKVELPSEQV 120
QY 365 GDLEASSEKCAPLATEVFDEKMEAHQ-EVVAEVHVSTVEKTEEQGGGGAEGGVVVEG 423
DB 121 SGSGPSEKPAFLATEVFDEKIEVHQEEVVAEVHVSTVEERTEEQ-----KTEVEE 172
QY 424 TGSLPPEKLAEPQEVQPAEAPAEELMKSRMCMVSGDHTQLTDLSPPEKTLPKHPGIV 483
DB 173 TAGSVPAEELVGMDAEPQEAEPKELVKLKTCTVSGEDPTQGDLSPEKVLSPPEGV 232
QY 484 SEVEMLSQERIKVQSGPLKLFSSGLKLSGKKQKGGGGDEPGEYQHHTESPE 543
DB 233 SEVEMLSQERIKVQSGPLKLFSTGLKLSGKKQKGR-KGGDEBSGHTQVPADSPD 291
QY 544 S 544
DB 292 S 292

RESULT 5
US-08-994-570-3
Sequence 3, Application US/08994570
Patent No. 6090929
GENERAL INFORMATION:
APPLICANT: Scott, John D.,
APPLICANT: Nauert, Brian J.,
APPLICANT: Klauck, Theresa M.
TITLE OF INVENTION: Protein Binding Domains of Gravin
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower/233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/994,570
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: No. 6090929and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 27866/33451
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 292 amino acids
TYPE: amino acid
STRANDEDNESS: single

TOPOLOGY: linear
US-08-994-570-3

Query Match 11.8%; Score 952.5; DB 3; Length 292;
Best Local Similarity 66.1%; Pred. No. 2.5e-51;
Matches 199; Conservative 30; Mismatches 53; Indels 19; Gaps 4;

QY 254 KDGEEKQKEPTKSPSPVNSSETTSFKFTHGWAGWRKKTSPKSKEDDLTAE 313
DB 1 KEEGEEKQKEPSKSAESPTSPVTSSTGTFKFFTCQWAGWRKKTSPKSKEDDEVEASE 60
QY 314 KRKEQAEKVDEEKEKTEPASEE-----QEPADTDQARLSADYKVELPLEDOV 364
DB 61 KKEQEPKVDTEEDGKAESAELKATASEQAHPQEPASAEHPRLSAEYKVELPSEQV 120
QY 365 GDLEASSEKCAPLATEVFDEKMEAHQ-EVVAEVHVSTVEKTEEQGGGGAEGGVVVEG 423
DB 121 SGSGPSEKPAFLATEVFDEKIEVHQEEVVAEVHVSTVEERTEEQ-----KTEVEE 172
QY 424 TGSLPPEKLAEPQEVQPAEAPAEELMKSRMCMVSGDHTQLTDLSPPEKTLPKHPGIV 483
DB 173 TAGSVPAEELVGMDAEPQEAEPKELVKLKTCTVSGEDPTQGDLSPEKVLSPPEGV 232
QY 484 SEVEMLSQERIKVQSGPLKLFSSGLKLSGKKQKGGGGDEPGEYQHHTESPE 543
DB 233 SEVEMLSQERIKVQSGPLKLFSTGLKLSGKKQKGR-KGGDEBSGHTQVPADSPD 291
QY 544 S 544
DB 292 S 292

RESULT 6
US-08-769-309A-17
Sequence 17, Application US/08769309A
Patent No. 5741890
GENERAL INFORMATION:
APPLICANT: Scott, John D.,
APPLICANT: Nauert, Brian J.,
APPLICANT: Klauck, Theresa M.
TITLE OF INVENTION: Protein Binding Domains of Gravin
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower/233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/769,309A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 5741890and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 27866/33451
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 651 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-769-309A-17

Query March	8.1%;	Score 656;	DB 1;	Length 651;
Best Local Similarity	38.3%;	Pred. No. 1.4e-32;		
Matches 191;	Conservative 80;	Mismatches 188;	Indels 40;	Gaps 17;

QY	1110	VANUV-----IKLOOL-MEOAARPSSEFLDSENGSTPLABSDTADGQOEDTDSQ	1163
QY	1110	VANUV-----IKLOOL-MEOAARPSSEFLDSENGSTPLABSDTADGQOEDTDSQ	1163
Db	6	VTTQOAEFTLAGVSGQEWMEQALPPBSVETPTDSETDGSTPLADDAAGTTOKEIBEIYEH	65
QY	1164	DSKATAVROSQUTEEBAATAQKEBESTLPNNVPAOEHEGEBPG-RDYLEPTOELTAA	1222
Db	66	BENEVHLV-PVRGTEAFAVPAQKERPPA-PSSFPQEBTKESQKMEDTLHHTDKESVET	123
QY	1223	VPLAKTEVGOEGBEVDMLDEKVK---EBQEVFVHSG-ENSQKADVTYDSEWGV	1275
Db	124	VSLSTKEGTQ--EADQYADEKTKDVPFEGLEGSIDTQITVSRKEVTEVALKQGEHEA	181
QY	1276	GCOKESTEVQS-----LSIEBEMETDYKEKEKTEPKQVSEEG-EDETAPEHEGTG	1329
Db	182	ECKDDALELQSHAKSPSPSEVEREMVVOVEREKTEAPHVAEKLEHHTAATVSEEVSK	241
QY	1330	KPVLTLDMPSEBSEKGLSGSGS-PSLPBDQKAGCIEVONOSLDTTVQTAFAVXVI-	1386
Db	242	QLQTVAVPPIIDGAKEVSSLEGSPPCLOGEAEVCTKQIVOSSEASFLLTAAEBEKVLG	301
QY	1387	ETVVISSESGSPCEVGAHLPAEKSSATGHWTLQHAEDTVLGPESQAESIPILVTPAP	1446
Db	302	ETANILETGETLEPAGHVLVEKSSSKMEDPAHNGEDAVPTGPDQAKSPVIVASTT	361
QY	1447	ESTLHPDLOGEISASQHERSEEBDKPDAGBDAGKESTALIEKVLAEPE-ILLESKSN	1504
Db	362	KKGLSSDLEGKTKTSLMKKSDDEVDEQAQCEV--KVSVAIEDP---EPENGILELETSS	416
QY	1505	KIVLVAVIQTAVDQAPAR-ETAPETHAVDSQTVQPAKRLDSREBNRCWK-----MKDKM	1558
Db	417	KLVOINIOTAVDQVPRTEETATENTJSELTOAHVIKADSOAGQETEKEGEEPQASQD	476
QY	1559	KHPVQPREDLQVLTVLEA	1577
Db	477	ETPITSASESESTAVGQA	495

RESULT 7
 US-08-994-570-17
 Sequence 17, Application US/0894570
 Patent No. 6090929
 GENERAL INFORMATION:
 APPLICANT: Scott, John D.,
 APPLICANT: Nautert, Brian J.,
 APPLICANT: Klauk, Theresa M.
 TITLE OF INVENTION: Protein Binding Domains of Gravin
 NUMBER OF SEQUENCES: 24
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 STREET: 6300 Sears Tower/233 South Wacker Drive
 CITY: Chicago
 STATE: Illinois
 COUNTRY: United States of America
 ZIP: 60606-6402
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/994,570
 FILING DATE:
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: No. 6090929and, Greta E.
 REGISTRATION NUMBER: 35,302
 REFERENCE/DOCKET NUMBER: 27866/33451
 TELECOMMUNICATION INFORMATION:

```

? TELEPHONE: 312-474-6300
? TELEFAX: 312-474-0448
? TELEX: 25-3856
? INFORMATION FOR SEQ ID NO: 17:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 651 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
?
?
US-08-994-570-17

```

Query Match	8.1%	Score 656	DB 33	Length 651
Best Local Similarity	38.3%	Pred. No. 1.4e-32		
Matches 191	Conservative	80	Mismatches 188	Indels 40
				Gaps
QY 1110 VATQV-----IKLOQ-LEQAVAPESSELTLDSENGSTPLADSDTAAGTQODETIDSQ 1163				
Db 6 VTTQAGATLTAAGVSGEMWEGAIIPDSVETPTDSETDSTPADPADGATGTXDEIVEIH 65				
QY 1164 DSKATAAVROSQVTEEDAAITQKEBPSTLPNNVPAQEEHGEBC- RDVLEPTQOELTAA 1222				
Db 66 EENEVHLV-PVRGTEAEAVPAQKEREPA-PSSFQFEQTEKESKMEDTLEHTDKEVSJET 123				
QY 1223 VPVLATKEVGOEDEVUMDGEKKY-----EBOEVPVSH--PNSKADAVDYDSEWGV 1275				
Db 124 VSLSTKEGTQ--EADQVADKTKDVPPFEGSGSIDGTIVSRKETEVALKSGTBEA 181				
QY 1276 GCOEKESTEVOS-----LSLEGEMETDYEKEREKTEKPEOVSEB-BOETAAPHEGTGY 1359				
Db 182 ECKDDALEIQSHAKSPSPSVYEREMVQYERKTEAPTHNNEKLEHETAVTVSEVSK 241				
QY 1330 KPVLLTDMPSSEKALGSLGGS-PSLPQDPAAGCTLENOVOSLDITVTQTAEVAKYI-- 1386				
Db 242 QLTQTVAVPPIIDGAKEVSSLEGGPPCQLQOEBAVCTKIQOVSEKSFLLTAAEKEKLG 301				
QY 1387 ETVVISETSGSPCEVGNHLLPAEKSSATGAGMHTLQHAEDTTLGPEQSAESIPIITPAP 1446				
Db 302 ETANILLETGETLEPAQHVLVEKSEKNEKEDPAHNGEAVYGTGDCQAKSTPIVSAIT 361				
QY 1447 ESTLHPDLQGEISASQRESESEEDKDPADGADGKESTALIEKVLKAPPE--ILELESKN 1504				
Db 362 KKGSLSDLEGKTKTSLKMKSCDEVDQVACQEV--KVSVAIEDL--EPENGILLELETKSS 416				
QY 1505 KIVLANVLOTAVDQAKI-ETAPETHAVDSQTVPRACRLDSRPNNCWK-----KKQAKM 1558				
Db 417 KLVONIITQTAVDQFVRTEETATMTLSQLOAHVIAKIDQADQGETEKEGEPQASQD 476				
QY 1559 KHPVQRPREDLOLVTLVEA 1577				
Db 477 EPTLSAKESBESTAVQQA 495				

RESULT 8
 US-08-769-309A-14
 : Sequence 14, Application US/08769309A
 : Patent No. 5741890
 : GENERAL INFORMATION:
 : APPLICANT: Scott, John D.,
 : APPLICANT: Nauert, Brian J.,
 : APPLICANT: Klauck, Theresa M.
 : TITLE OF INVENTION: Protein Binding Domains of Gravin
 : NUMBER OF SEQUENCES: 24
 : CORRESPONDENCE ADDRESS:
 : ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
 : STREET: 6300 Sears Tower/233 South Wacker Drive
 : CITY: Chicago
 : STATE: Illinois
 : COUNTRY: United States of America
 : ZIP: 60606-6402
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/769,309A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 5741890and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 27866/33451
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 453 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-769-309A-14

Query Match 7.9%; Score 639.5; DB 1; Length 453;
Best Local Similarity 40.5%; Pred. No. 8.7e-32;
Matches 183; Conservative 69; Mismatches 165; Indels 35; Gaps 16;

QY 1110 VATCOV-----IKLOQL-MEQAVAPESSETLTDSETNGSTPLADSDTADGTQOQDETIDSQ 1163
Db 6 VTTCOAETLAGVKSQEMVMEQAIPDPSTVETDSETDGTVPVADFPDAGTGTQKDEIVEIH 65
QY 1164 DSKATAAVRSQVSTEEAATAQKEEPTLNNVPAQEEHGEPEG-RDVLEPTQOELTAA 1222
Db 66 ENEVHLV-PVRGTEAEVPAQKERPPA-PSSFVFQETKEQSKMEDTLEHTDKESVET 123
QY 1223 VPVLAKEVTEQGEVDWLDGEKVK-----EEQEVFVHSG--PNSQKAADVTYDSEVMGVA 1275
Db 124 VSILSKTEGTQ--EADQYADEKTKDVFFEGLEGGIDGITVSRKVTVALKGEETEA 181
QY 1276 GCOEKESTEVQS-----LSLEGEEMTDVEKRETKPQOVSEEG-EQETAAPEHEGTYG 1329
Db 182 ECKKDDALELQSHAKSPSPVEREMVQVEREATEPETHVNEEKLEHETAVTVSEVSK 241
QY 1330 KPVLTLDMPSSERKALGSLGGS-PSLPDODKAGCIEVQVQSLDITVTTQTAABAEVKVI-- 1386
Db 242 QLLQTVNVPIDGAKVSSLEGPPCLGQEEAVCTKIQQVSSASLTTLTAAABEEKVLG 301
QY 1387 ETVVISGTGSPCEVGAHLPAEKSSATGCHWTLQHAEDTVPLGPESQAESIPITVPAP 1446
Db 302 ETANILETGTLEPAGAHVLVEKSSKEDFAAHPCGEDAVPTGPDCAKSTPVIYSATT 361
QY 1447 ESTLHPDLOGEISASQRESEEDKPDAGPDAGKESSTAIEKVLKAEPE--ILELESKSN 1504
Db 362 KGLSSDLEGEKTTSLKWKSDVEVDEQVACQEV--KVSVAIEDL---EPNGILELETKSS 416
QY 1505 KIVLNVIOQTAVDQFART-ETAPETHAYDSQTQ 1535
Db 417 KLVQNIIOQTAVDQFVRTEETATEMLTSELQTQ 448

RESULT 9
US-08-994-570-14
Sequence 14, Application US/08994570
Patent No. 6090929
GENERAL INFORMATION:
APPLICANT: Scott, John D.,
APPLICANT: Nauert, Brian J.,
APPLICANT: Klauck, Theresa M.,
TITLE OF INVENTION: Protein Binding Domains of Gravin
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower/233 South Wacker Drive
CITY: Chicago
STATE: Illinois

COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/994,570
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: No. 6090929and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 27866/33451
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 453 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-994-570-14

Query Match 7.9%; Score 639.5; DB 3; Length 453;
Best Local Similarity 40.5%; Pred. No. 8.7e-32;
Matches 183; Conservative 69; Mismatches 165; Indels 35; Gaps 16;
QY 1110 VATCOV-----IKLOQL-MEQAVAPESSETLTDSETNGSTPLADSDTADGTQOQDETIDSQ 1163
Db 6 VTTCOAETLAGVKSQEMVMEQAIPDPSTVETDSETDGTVPVADFPDAGTGTQKDEIVEIH 65
QY 1164 DSKATAAVRSQVSTEEAATAQKEEPTLNNVPAQEEHGEPEG-RDVLEPTQOELTAA 1222
Db 66 ENEVHLV-PVRGTEAEVPAQKERPPA-PSSFVFQETKEQSKMEDTLEHTDKESVET 123
QY 1223 VPVLAKEVTEQGEVDWLDGEKVK-----EEQEVFVHSG--PNSQKAADVTYDSEVMGVA 1275
Db 124 VSILSKTEGTQ--EADQYADEKTKDVFFEGLEGGIDGITVSRKVTVALKGEETEA 181
QY 1276 GCOEKESTEVQS-----LSLEGEEMTDVEKRETKPQOVSEEG-EQETAAPEHEGTYG 1329
Db 182 ECKKDDALELQSHAKSPSPVEREMVQVEREATEPETHVNEEKLEHETAVTVSEVSK 241
QY 1330 KPVLTLDMPSSERKALGSLGGS-PSLPDODKAGCIEVQVQSLDITVTTQTAABAEVKVI-- 1386
Db 242 QLLQTVNVPIDGAKVSSLEGPPCLGQEEAVCTKIQQVSSASLTTLTAAABEEKVLG 301
QY 1387 ETVVISGTGSPCEVGAHLPAEKSSATGCHWTLQHAEDTVPLGPESQAESIPITVPAP 1446
Db 302 ETANILETGTLEPAGAHVLVEKSSKEDFAAHPCGEDAVPTGPDCAKSTPVIYSATT 361
QY 1447 ESTLHPDLOGEISASQRESEEDKPDAGPDAGKESSTAIEKVLKAEPE--ILELESKSN 1504
Db 362 KGLSSDLEGEKTTSLKWKSDVEVDEQVACQEV--KVSVAIEDL---EPNGILELETKSS 416
QY 1505 KIVLNVIOQTAVDQFART-ETAPETHAYDSQTQ 1535
Db 417 KLVQNIIOQTAVDQFVRTEETATEMLTSELQTQ 448

RESULT 10
US-08-769-309A-15
Sequence 15, Application US/08769309A
Patent No. 5741890
GENERAL INFORMATION:
APPLICANT: Scott, John D.,
APPLICANT: Nauert, Brian J.,
APPLICANT: Klauck, Theresa M.,
TITLE OF INVENTION: Protein Binding Domains of Gravin

```

NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESSES:
ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower/233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/769,309A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 5741890and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 27866/33451
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 396 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-769-309A-15

Query Match      6.5%; Score 521.5; DB 1; Length 396;
Best Local Similarity 38.2%; Pred. No. 1.3e-24;
Matches 147; Conservative 62; Mismatches 149; Indels 27; Gaps 12;

QY 1110 VATCV-----IKLOOL-MEOAVAPESSETLTDTSETNGSTPLADSDPTADGTQODETTDSQ 1163
D 6 VTTQAEETLAGVKSQEMWMEQAIPPDSVETPTDSETDGTVPADFDPACTQKDEIVEIH 65
QY 1164 DSKATAARQSQVTEBEATQKEEPTLPNNVPAQEHGEEPG-RDYLEPTQOELTAA 1222
D 66 EENEVHLV-PVKGTEAEAVPAQKERPPA-PSSFVFOETKEQSKMEDTLEHTDKEVSVE 123
QY 1223 VPVLAKTVEGGEVDMLDGEKVK-----EEOEVFVHSG--PNSOKAADVTYDSEWGV 1275
D 124 VSIISKTEGTQ--EADQYADEKTDVPPFBGLSGSIDGTITVSREKVEVALKGEETEA 181
QY 1276 GCOEKESTEVQS-----LSLEGEMETDVEKEKRETPREQVSEEG-EOETAPAEHGTYG 1329
D 182 ECKKDDALELQSHAKSPSPVEREMVQVEREKTEAEPTHVNEEKLHEHTAVTVEEVS 241
QY 1330 KPVLTLMPPSSERKALGSLGS--PSLPDQKACIEVQVOSLDTTVOATAVVKVI-- 1386
D 242 QLLQTVNPIIDGAEVSSLESGPPCLGGEAVCTKIQVSSSEASFTLTAABEEKVLG 301
QY 1387 ETVVISEGESPECVGAHLPAEKSSATGHWTLQHAEDTVPLGPESQASIPITVPAP 1446
D 302 ETANLLETGETLEPAGAHVLEEKSEKNEDEPAHPGDAVTPGDCQAKSPVIVSAT 361
QY 1447 ESTLHPDLOGEISASQERSEEDK 1471
D 362 KKGSSDLEGEKTSLKWKSDEVDE 386

```

```

APPLICANT: Klauk, Theresa M.
TITLE OF INVENTION: Protein Binding Domains of Gravin
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESSES:
ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower/233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/994,570
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: No. 6090929and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 27866/33451
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 396 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-994-570-15

Query Match      6.5%; Score 521.5; DB 3; Length 396;
Best Local Similarity 38.2%; Pred. No. 1.3e-24;
Matches 147; Conservative 62; Mismatches 149; Indels 27; Gaps 12;

QY 1110 VATCV-----IKLOOL-MEOAVAPESSETLTDTSETNGSTPLADSDPTADGTQODETTDSQ 1163
D 6 VTTQAEETLAGVKSQEMWMEQAIPPDSVETPTDSETDGTVPADFDPACTQKDEIVEIH 65
QY 1164 DSKATAARQSQVTEBEATQKEEPTLPNNVPAQEHGEEPG-RDYLEPTQOELTAA 1222
D 66 EENEVHLV-PVKGTEAEAVPAQKERPPA-PSSFVFOETKEQSKMEDTLEHTDKEVSVE 123
QY 1223 VPVLAKTVEGGEVDMLDGEKVK-----EEOEVFVHSG--PNSOKAADVTYDSEWGV 1275
D 124 VSIISKTEGTQ--EADQYADEKTDVPPFBGLSGSIDGTITVSREKVEVALKGEETEA 181
QY 1276 GCOEKESTEVQS-----LSLEGEMETDVEKEKRETPREQVSEEG-EOETAPAEHGTYG 1329
D 182 ECKKDDALELQSHAKSPSPVEREMVQVEREKTEAEPTHVNEEKLHEHTAVTVEEVS 241
QY 1330 KPVLTLMPPSSERKALGSLGS--PSLPDQKACIEVQVOSLDTTVOATAVVKVI-- 1386
D 242 QLLQTVNPIIDGAEVSSLESGPPCLGGEAVCTKIQVSSSEASFTLTAABEEKVLG 301
QY 1387 ETVVISEGESPECVGAHLPAEKSSATGHWTLQHAEDTVPLGPESQASIPITVPAP 1446
D 302 ETANLLETGETLEPAGAHVLEEKSEKNEDEPAHPGDAVTPGDCQAKSPVIVSAT 361
QY 1447 ESTLHPDLOGEISASQERSEEDK 1471
D 362 KKGSSDLEGEKTSLKWKSDEVDE 386

```

```

RESULT 11
US-08-994-570-15
Sequence 15, Application US/08994570
Patent No. 6090929
GENERAL INFORMATION:
APPLICANT: Scott, John D.,
APPLICANT: Naureit, Brian J.,

```

```

RESULT 12
US-08-973-462-8
Sequence 8, Application US/08973462B
Patent No. 6191270
GENERAL INFORMATION:

```

RESULT 13
US-07-853-913-2
; Sequence 2, Application US/07853913
; Patent No. 5338839
; GENERAL INFORMATION:

```

APPLICANT: McKay, Ronald D.G.
TITLE OF INVENTION: Nestin Expression As An Indicator of
TITLE OF INVENTION: Neuroepithelial Tumors
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESS: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/853,913
FILING DATE: 19920319
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/660,412
FILING DATE: 22-FEB-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/603,803
FILING DATE: 25-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/201,762
FILING DATE: 02-JUN-1988
PRIOR APPLICATION DATA:
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: MIT-4641AAA
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1805 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-853-913-2

Query Match 4.8%; Score 388; DB 1; Length 1805;
Best Local Similarity 21.3%; Pred No. 1.8e-15;
Matches 319; Conservative 233; Mismatches 543; Indels 402; Gaps 75;

QY 13 EOPAGSDT---PSELV-----LSGHAPAAEASGADPADADPATKLPOKNGQLSSVNGV 64
DB 440 EEPGQGOQHFPDDLTSLATNLNPHNPTLEAKO--GESESRSVSIPODEQIWL--V 495
QY 65 AAGQDVHVOENQEQ--EEVVDDEVQGESDVAEKRVEMMAN---STRAVEDITMD 119
DB 496 EKEADIEVENSASAKTQESGLDTEQ--DSQGPLQKETLKLAGEPLMSLKIQYETA 554
QY 120 GGE-----ETSEIITQIP---ASENNVEEMVQAPASSQANDGVFKVFPFVG 163
DB 555 GKENCSSTEGHGLTLEGPKEKQIPKLSEKQVE-----SEKTLGVPLVLSLLGK 608
QY 164 KETVKKDKNEKSDTVQLTVKKDEGEAGASVAGDHOEPSVET-----AVGESA 213
DB 609 EDTRTEDQELMSKGS---TLKRSSLSGKESQEVVVRPSKGNLSWAFKESQHPGLGPG 665
QY 214 SKESSELKQSTEQEGTL---KQESSTELPLQESDQ--AAEEADDEGE-----EKGE 262
DB 666 AADQMLERLVEKEDQSPSPSEEDQACRPIQKENQEPGLGYEEA--EQQIERLLTKES 723
QY 263 KPTSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 322

```

```

DB 724 QESLRSPSEB---DQAGRSLOK-----ENQEPVGYEAEADQMLERLTKESQESLK 772
QY 323 VDEEKETKTPASHEQ-----PAEDDQALSA-----DYKVELPLE-- 361
DB 773 SPENORIGPLERENQSLRYLEENQETVPLESNRORPLRSLEVEEORIVKPLEKV 832
QY 362 --DQVGLASSEKCAPLATEVDEKMAHQEVAHVHSTVEKTEEEOGGGGAEGV 419
DB 833 SQDSLGL---AEENVPPLR---YLEEDDCINKSL-----DKTKLSLSLEDRNGDS 880
QY 420 VV-----EGTGESL--PPEK-----LAEPQVQEAEPAPAEIMKSRKCVSGGDHTOL 465
DB 881 IIPQSEETQVSLRPPSEEDQRIYNHLEKESQEFSSSEEEQVM---ERSLEGENHSL 937
QY 466 TDLSP-----EKLTPKPEGVSEVEMLSQERIKVQGSPLKLFSSSGIKLSGKKQK 521
DB 938 SSVKEQOWVESQLEKSSQSGKSLD--EQETP---GPLEK--ENAESRSLSLGOQ-- 989
QY 522 KRGGGDEEPGEYOHITHESPESADQKGSASSASPEPEPTTCLKGPLAPQD--GEAE 580
DB 990 -----EQKLPQETQTLRAVNGMAVSP-----EKVDEPLPKPLGNDQ 1030
QY 581 EGTSDKEKKREGITPWASFKQWTPPK--RVRRPESDKEEPLKYKS---ATLSST 633
DB 1031 EIASLQKENQESLVSLEK--KGIEYVKSLEIEIIEPLET--AEEDLERKKSIDTQEPIMST 1088
QY 634 D---STYSEMDP-----VKTYGEOKP-----EPPKRVDTSVSWEALICVSSKKRARK 682
DB 1089 EVARETYPPPEDEPPGSLGVDENRETLSLEKSSQELSSGKNNVETRVEDSQCLQVE 1148
QY 683 SSSDEGGPRTL-----GDSHRAEASKDEAGTDVAPAST----- 719
DB 1149 EGLQEOHQESLREVQKQLPSSGNOQRMEDVBEKAVQGAHPLATTGTEDKAELHNG 1208
QY 720 --QEQDAQ-----GSSSPAPAGSPSBC---EGVSTWESFKRLVTPPKK 758
DB 1209 QGGEBAABEGELLQDIYGAWSLSGSEPKQVRPAEALDNLLEGGA-----LEVPAVQ 1261
QY 759 SKSLKEKAD--SSVRLSTIEIPSRRESVSIKKP1PGRKKRAGCKOQATVEDSGP 816
DB 1262 SMPVETERDEEDRAQAGQDSLEVTGLGAARTGLEL-----EQGVGLSD--P 1307
QY 817 VEINDDPNPAVPLSEYNVVEREKMAQNTLPLQALVAVYS--BELSTLVHTVSA 875
DB 1308 RHFAREKALPPS---LGEESKAKIAGLEBPGKEPKENAGLBDGILELPTTSSALECQ 1364
QY 876 VIDGRAVTSVEERSPWSISAVTEPLEHTAGAMPVVEVETKDIIAET----- 926
DB 1365 GHEEESMEGEEEBAS-----LETSDBEGSDAPQRPPEETDEGAQALGTAPKPLL 1418
QY 927 -----PVLTOPLRGKADHDMWTSE-----VPTSAVAVATETSEALRTBEVTEA 972
DB 1419 EPCSPILPLT-----DAHLQPAEGIQEAGWQPEXGSSALRVERPEFGAGEIPEG 1471
QY 973 -----SGAEETDMVASVQLTSPDTTEEATP---VOEVEGVLDTREBEROTALLQ 1023
DB 1472 LQDMWEGREB-----SEADDLQETLPDSTPLGLYRASASPKWDLAGOR-----LS 1518
QY 1024 AVADRYKKE--SQVPATQVORTGSKALEKVEVEEDEVLASEX-----EKDVMKGP 1075
DB 1519 PGDAGKEKEDWGPVAPAAQ-----GLSGPREEBEGHGSJLSEFEFDLGTASLIPVP 1573
QY 1076 VOEAQAEHLAQ-----GSEFGQATPREL--EVEEVTAADVHVAT----- 1112
DB 1574 KEV--ADHVQVPPVLPACMDQGESDGFADBEESGEEGEEDDEGASGAGQMWSSG 1631
QY 1113 -----COYIKLQQLMEQAVAPSSSETLTDSETNGST-----PLADSPDADQTODETID 1161
DB 1632 ASGGCKVQODIAQRDPPVOESVGVSGLWDDGLRGAANAANVPALEWVSQSAEBSGSEES-- 1689
QY 1162 SODSKATAVROSQVTEEEAATAQKEBPSTLPNNVPAOEHEGEBGRDVLBEPTQOELTAA 1221
DB 1690 ---ESASLEGEQGVTDH--LDAPQEVTSWVPGVGDADIDIGQSPNLD----- 1732

```

Qy 1222 AVPLAKTEVG---QEEDVMDLGEKVEEVEF-----VHSGPN 1258
Db 1733 SEQVNGRMENLEQAEQGV-VLDGD---EDQELLQGEVGAUKVPLVASPVHLGPS 1785

RESULT 14
US-09-134-001C-5080
; Sequence 5080, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5080
; LENGTH: 3696
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5080

Query Match 4.5%; Score 365.5; DB 4; Length 3696;
Best Local Similarity 18.2%; Pred. No. 1.2e-13;
Matches 324; Conservative 309; Mismatches 694; Indels 449; Gaps 75;

Qy 46 DPATKLPQKQGL-----SSVNGVAEQGDVHVQ-----EENQGEVEEVDVQGESDV 97
Db 1462 DAITSLSKINAQVVEKESARNAIEQAKTQQTQFINNDNDATDEEKEVANNILVIATKQKSL 1521

Qy 98 REKDRVEMAANSTAVEDITKQGESEIIEQIPASENNVVEVQPAESQANDVGPKV 157
Db 1522 ---DNINSLSNND- VENAUVAGINEANVLPAVAVSKAKKIDQKLAQINQIOTHT 1577

Qy 158 FKFGVFTVKDKNEKSDTVOLLTVKKDEGEAEASVGAGDHQEPSVETA----- 208
Db 1578 -----ATTEKEAAIQLANQKSNERTAIQN-----EHNNGVQAQKNGIHEIE 1622

Qy 209 -VGSASKESELKOSTKEQGTLLKQESSTIPIQAESDQAAE-----AKDEGE 258
Db 1623 LVMPDAHKKSDAKQSIDNK---VNEQSN---INTTPDATDEEKQKALDKLAKDAGY 1675

Qy 259 EKQEKPTKSPSPSPNSPSETTSFKFFTHGWAGWRKTSFK---KSKEDDLE----- 310
Db 1676 NKVDQAOTNQOVSDAKTEAIDTITNIQ-----ANVAKFSARVELDSKPEDLKQINA 1728

Qy 311 --TAEKKEQEA-----EKVDEEKEKTEPASEE-----PAEDTD- 345
Db 1729 TPNATEEKQDAIQLRNGKRDENVKLNINQRRDNEVEQHKNIQLELTHANTRKSDA 1788

Qy 346 ---QARLSADYKVELPLEDQVGDLEASSEK-----CAPLATEV 382
Db 1789 LQELQTFISQTELIN-----NNKDATNEEKDEAKELLBSKKNITINQATNNQV 1841

Qy 383 FDEKMEAHQEVVAEVHVHTVEKTEBEGGGEAGGV-VWEGTESILPPEKLAEPQEVQ 441
Db 1842 DNAKDNMGNEIATIPATTI-KTDAKTAIDKKAQEQVVTIINGNNDATDEEK-ABEARKLVE 1899

Qy 442 EA--EPAEELMKS-REMCVSGDHTQJTDL-----SPEEKTLPKHPEGIYSEVEMLSQE 493
Db 1900 KAKTEAKSNITNSTEREVNGAKTNGLEKINNIQFSTQTKTNAR-----QEINDKAEQ 1953

Qy 494 RIKVQSGPLKKLFSSSGLKSLGKQKRGKGDDGDEPGEYQHI-----HTESPESADEQK 549
Db 1954 LIQINNTP-----DATEEEKQATNRVAGLAQAQININNAHSTQEVNEK 1999

Qy 550 GESSAS-----SPEEPETTCLEKGLSAPQDGE--ABEGTTSDEKKREGITPWASPKKM 603

Db 2000 TNSIATIKSVQPNVIRKPTAINSITQEANNQKTLIGNDGNATDDEKE-----AAQOL 2051
Qy 604 VTPK--KRVRPSPESDKKEBELEKVSATLSS---TSTVSEMDEKVTGGE---EQKPEEP 656
Db 2052 VTQKLEQIQIKIHESIQDNQVNVKAQAITAIKLINANAHRQDAINILNLAESKSDI 2111
Qy 657 KRRVDTSVSWE--ALICVGSKKRARKASSDDGGGRTLGGDSHRA--BEASKDKKAGT 712
Db 2112 RANQDATTEEKNTAIOSIDDTLAQAR-----NNINGANTNALVDLENLEDGKQL 2160
Qy 713 DAVPASTQEODQAGSSP---EPAGSPSGEGVST---WESFKRL----- 752
Db 2161 QRVLSTQTTQAKADIAQIGQORSTIDQONATTEEQEALERLQNETNGVNDRIQAA 2220
Qy 753 -----VTPRKKS-----KSKLEEKAEDSSVEQ---LSTEIPSPSEESVSI 790
Db 2221 LANQNVTDENKNNILETIRNVEPIVVKPANEIRKKAQEQTLINQONQATLEEKQIAL 2280
Qy 791 KKFIPGRK-----KRADGKQEQATVEDSGPVEINEDDPNVPAVPLSEYNAREKMEA 845
Db 2281 GKLEEVKNEALNQVSAHNSNDVKIAENNGIAKISEVHPET-----IIKENA--KQIEQ 2333
Qy 846 QGNTELPQLLGAVVVSBEELSKTLVHTVSVAVIDGTRAV-----TSVEE 888
Db 2334 DAOSQIDTINANKSTNEEKSAAIDRVNVAIDAINNITATTQTOLVNDKNGSNTSISQ 2393
Qy 889 RSPSWISASVTEPLEHTAGEMPVEEVEKDIIEAETPVLTQTLPGKDAHDMVTSEV 948
Db 2394 ILPS--TAVKTNALALASA-----KKNKAIIDOTPNATAEKEEANKVDRLOEEA 2444
Qy 949 DFTSEAVTATETSEALRTEEVTEASGAE-----ETTDMVSAVSQTLDSPTTEATPVQEV 1004
Db 2445 DANILKAHTTDEVNNIKQAVQINAVQEVIKQNVQNLQNFIDNOKKIIENTPATL 2504
Qy 1005 ESGVLTDEEERQTOALQAVADKVKESQVPATQTVORTGSKALEKVEVEEDSEVLAS 1064
Db 2505 E-----EKAEANRLQNLVLTSTSEI---ANVDHNNVEDQALDKARPKIEE---VPQV 2552
Qy 1065 EKEKDVMPKPGVQAGAEHLAQSGEQATPESPESLEVPVTDVHDVATCOVILKQOLMEQ 1124
Db 2553 SKRDVL--NAIQEAFNSQTEIQENQEATNEE-----KTEALKNQLLNQ 2597
Qy 1125 AVAPESSETLTDSETNGSTPLADSTADGTQDDETIDSQDSKATAAVRQSOVTEE----- 1179
Db 2598 AKV-----NIDQASNKVDSDAKTSRSIQDIEQIQPHQPTKATG 2635
Qy 1180 -----EAATAQKSEPTLPNNVPAQEBEGEPGRDVLPTQOELTAAAVPVLAKEVGOE 1234
Db 2636 RHLNEKANQQOSTIATHPNSTTEERQESAKLOEVLK-----KAIKIDKQGT 2684
Qy 1235 GEVDWLDGK-----VKEQEVFVHSGPNSQKAADVTDSEVMGVAGCQSESTEVSLS 1289
Db 2685 ND-----DVEKTVNGTAIEINILPATVTKDKAKADV-----AEKQKNLQINS 2729
Qy 1290 LEEGEMETDVEKRETKPEQVSEEGEQUETA-APHEGTYGKPVLTLDMPSSERGAALGS 1348
Db 2730 NDEATTE---EKLVASDNLNHHVETTQALDAPD-----TNQNVNKKNGIGT 2775
Qy 1349 LGG-SPSLPDQDKA-GCIEVQVQSLOTTVTQTAEAV-----EKVIETVVI 1391
Db 2776 IRDIQPLVVKKPTAKGIESAVEKKKTEINQTNATHDEVREGLNQLNQHAKAKDNVQ 2835
Qy 1392 SETGESPECVGAHLLPAEKSSATGCHWTLQHAEDTVPGLPESQAESIPITVTPAPESTLH 1451
Db 2836 SQTNQVVE-----NAEQNS-----LDQINNFRP--DFSKRNAAVEIVKAQONKI- 2878
Qy 1452 PDLQGISASQBRERS-----EBE-----DKPAGPADGKESATAIEKVLKAEPEIL 1497
Db 2879 DEIEQFESATQEEKDNALQHLDOVKIEINSINQANTDNEVDNAKTSGLNITEYREPY- 2937
Qy 1498 ELBSKSNKIVLNIQTAVDQFARTAP-----ETHAVDSQTVQVPAKRLDSR----- 1544

Db 2938 ---KKKNAIILKYDVSDTGEAIIINGYPDATEDELOEANSKLINKILLDAKKOIGIAHTNN 2994
 Qy 1545 EPNRCWTKMKDAKKMHPVPOFPREDLOVLTYLEMAQ 1580
 | : : : || : : : || : : :
 Db 2995 EVDVDIYNEVSO-KMKTLIPRVDTKAAVAKRSVTLNLAQ 3029

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Query Match 4.5%; Score 361.5; DB 1; Length 2101;
Best Local Similarity 19.6%; Pred. No. 9,6e-14;
Matches 364; Conservative 292; Mismatches 668; Indels 529; Gaps 81;

QY	53	OKGQGLSVNVAEGDVHYVQENQ--EGQEEVVDVQVQRESDVREKDRV-----	103
Db	357	EKOQOLEKEKLSALQDDKKCCLEKEKILQKLSQ--LEBHSLOLQDNPRQEGVILQDVLQ	415
QY	104	-----EEMANSTAYED-----ITKQGOEET-----SEITEQIPASE	135
Db	416	ETLKQEAATLANNITQQLQARVEMLETERGQGOEAKLLAERGHFEEKQQLSLLITDQSSIT	475
QY	136	NNV-----EEMQVPAESQA-----NAGVFQKVPYFGFQFYKQDKNEK	174
Db	476	SNLSQAKKEELQASQAHGARIITAOVASITSELTLLNNTITQOQDQDLGLQQAQNEKQALQ	535
QY	175	SDTVQLITVKKDEGGAEGASVAGDHQEPVETAVGESASKESELKOSTQEGITLKQDQ	234
Db	536	AQTLQ-----QQEGASQGL-RHQ--VEQLSSILKQEGQLKVEAKQESAT--RQD	580
QY	235	STSEITPLQABSDQA--EEKADGEQEKEQEPKPSQ--SPSSPVNSITSSQFKFFTH	280
Db	581	HAQQLATAAEREERASLRDRDALQLELEKEKAKLEITLQOQLQVANEKRDASQVSTQ	640
QY	291	GWAGWRKRTTSFKSKED--DLETAKEKKEQDAEKVDEE--KEKTEPASQEDQEPADT	344

Dd		641	---	AQRKALSKVBELOQACVETARQOEHEQAQVALEIQLBSBOQAKTEKRVAOQ	697
Oy		345	DQARLSADYKEVELPLEDOVGDLNASSSECKAPLATEVPFERMEAOEVAEV----	H 398	
Dd		698	DQ--LOBOLQALKESILKVTKGSIE---EBK--RRAADALEBQRCISIELKRETSLEVQH	750	
Oy		399	VSTVEKTEBEOGG-----GGAEGGVNBEGTSGSLPEKLAERQ-----	437	
Dd		751	KRKERELEEBAARKGLBARLLQLGEAHOARETVLRRELAEMAAQHTABSECHOLVKEV	810	
Oy		438	-----EVOQEAEP-----AELMKSRMENVSGSDHQTOLDLSPBEKTLPRKHPEGIVS	484	
Dd		811	AAAMDGVEDBOQEAOYGAMFQBOLMTLKEBEEKAROELQ--EAENKV----AGIES	861	
Oy		485	EVENLSSQERIKNVQGSPLKLRS--SSGLKISGKKQKRGKGCGGDEEPGEYOHIHTESP	542	
Dd		862	HSELQISRQONK-----LAEIHANLARALOQVOEKVBQK--LADDLSTLOEKMAATSK	914	
Oy		543	ESADEQGESSASSPEBPETTCLF--KEPLA-----PODGAEBGTSSDEKKRBGITP	596	
Dd		915	EVA---BLETLVRKAGEQOETASRELVKRPADAGROPWELEOOG-----ROFCST	963	
Oy		597	WAFKKAVTPEKKAVERPSBEDKEE---ELEKYSATLSTDSITVSE---MODEVKTAGEE	650	
Dd		964	QALQAM-----BREAOGMENELERRALAMESQQGOOEEROQOREVFARLTOE	1012	
Oy		651	QKPEPKRVDTSVSWBALICVSSSKRRARKASSDDDEGCPRTJGSDSHRAEASKOE-	709	
Dd		1013	RGRQADIALFEKARAARELMRLQNALNEQRFVFPATLOELAHAL-----TEKGXQDEL	1066	
Oy		710	---AGTDAY-----PASTODODAQOSSSPERGSPSBEGSTWESF	749	
Dd		1067	AKLRGLEPAQITKELEBLROTVKOLKEQLAKKEHKHSNGAOSEAGTBETGPRL-BAL	1125	
Oy		750	KRLVTPRKKSXLKEKAE-----DSSYEOLSTEIPRSRESWSIKKF	793	
Dd		1126	RAEVSUKLEQOCQOQOEBQDLSLERLSAERBASARBDSALFTLOOLEKQACHEQSQAL	1185	
Oy		794	I PRRKKRA--DGKOEOATVEDSGPVET---NEDDPNVPAVPL-SETYNAVEREKMEAQ	846	
Dd		1186	ASAORELAAPRTKYODHSKSEDPMKQAVARGOEARNSLSTLSDEEVSITLNQVLKE	1245	
Oy		847	GNTLPOLLGAVVYSESLSTLVHTVASVANVDOTRAVTSVEERSPSWISASVTPELHTA	906	
Dd		1246	G--ESKEKLKWALESKSKULEBSQACC-----KORQOA-----	1278	
Oy		907	GEAMPVEEYTEKDIIAETPVLTOTLPBGKDAHDMWTSE---VDFTSEAVTATETSEA	963	
Dd		1279	-----TYPELQNALLCGR-----RCRAGREAEKQVASENLRLBELTSGEABELQOE	1328	
Oy		964	LRT--EEVTEASGAEEITDMVSAVSOLTSPDTTEBATPVQVESGVLDITEEROTQAI	1021	
Dd		1329	LKAWOEKFPOEQRALSTLOU-----EHSTQALVELLPAHNLCQOLOAE	1373	
Oy		1022	LQAVADYKKESSQVPAQOTQORTGSKALE-----KYEBVEDSEVLASE	1065	
Dd		1374	QAAAEKRRRELE---QSKQAAAGGLRAELLRAQRELBELRLRQKVAEOERTAOQLAAE	1429	
Oy		1066	K-----K-----BKDWMPKGPVOEAGAE	1082	
Dd		1430	KASYABOLSMUKAHGLIAENKGLBERANGLERNOFLVEILDQARBKXYOELAAYADAET	1489	
Oy		1083	HLAQSETFQATPSPLEVEPVTAADVHDVATCOVIKLOLMEOAVAPBESSELTITDETNGS	1142	
Dd		1490	RLAEVORAOSTARELEV--MTAKYEGA-----KKYVLEEQRFOEBEROKLTQOVELS	1541	
Oy		1143	TPLDSPDACTOOD--ETJDSC--OSKATTAARVQSVTEEAAATQAKEBESPTLPNNVPA	1198	
Dd		1542	KKLADSDQASVYQOQKAKAVQADGESQOEBAQFQAQLMBELOALSQYEQAA-----	1593	
Oy		1199	QOEHGEBGBRDVLEPTQOELTTAAAVPLKTEVGOEGBVMDLIGERVXKBEOEVPVHSGPN	1258	

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 13, 2002, 00:16:06 ; Search time 224 Seconds
(without alignments) · 118.719 Million cell updates/sec

Title: US-09-902-432-4
Perfect score: 8073
Sequence: 1 MCAGSSTEQRSPSEQPAGSDT.....AWAQRKCLPRLQLKAPVSK 1596

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105981 seqs, 16662342 residues

Total number of hits satisfying chosen parameters: 105981

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

```

Database : Published Applications AA: *
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pcp.*
2: /cgn2_6/ptodata/1/pubpaa/FCT_NEW_PUB.pcp.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pcp.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pcp.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pcp.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pcp.*
7: /cgn2_6/ptodata/1/pubpaa/FCTUS_PUBCOMB.pcp.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pcp.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pcp.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pcp.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pcp.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pcp.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pcp.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pcp.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	8073	100.0	1596	9	US-09-902-432-4	Sequence 4, Appli
2	5421.5	67.2	1346	9	US-09-902-432-2	Sequence 3, Appli
3	4294	53.2	1781	10	US-09-738-877-3	Sequence 3, Appli
4	462	5.7	1786	9	US-09-743-096-3	Sequence 3, Appli
5	429	5.3	2665	10	US-09-864-761-34248	Sequence 34248, A
6	407.5	5.0	2368	10	US-09-815-242-5635	Sequence 5635, Ap
7	407.5	5.0	2368	10	US-09-815-242-12389	Sequence 12389, A
8	388	4.8	1325	10	US-09-864-761-35612	Sequence 35612, A
9	382	4.7	1879	9	US-09-971-536-70	Sequence 70, Appl
10	377	4.7	2478	10	US-09-815-242-5816	Sequence 5816, Ap
11	377	4.7	2478	10	US-09-815-242-12967	Sequence 12967, A
12	377	4.7	3256	10	US-09-919-172-98	Sequence 98, Appl
13	368	4.6	1400	10	US-09-764-176-7	Sequence 7, Appli
14	354	4.4	1618	9	US-09-963-875-1	Sequence 1, Appli
15	351.5	4.4	617	10	US-09-864-761-36182	Sequence 36182, A
16	339.5	4.2	2139	10	US-09-737-384-6	Sequence 6, Appli
17	334.5	4.1	2344	10	US-09-815-242-12713	Sequence 12713, A
18	323.5	4.0	1884	10	US-09-785-770A-1	Sequence 17, Appl
19	323.5	4.0	1907	10	US-09-785-770A-16	Sequence 16, Appl

20	322	4.0	2310	9	US-09-991-496-120	Sequence 120, App
21	322	4.0	2310	10	US-09-874-923-120	Sequence 120, App
22	314	3.9	1863	9	US-09-734-672-4	Sequence 4, Appli
23	313.5	3.9	2843	9	US-09-987-482-1	Sequence 1, Appli
24	312.5	3.9	2843	8	US-08-681-219-32	Sequence 32, Appli
25	308.5	3.8	26926	9	US-09-759-508B-2	Sequence 2, Appli
26	306	3.8	1863	9	US-09-734-672-2	Sequence 2, Appli
27	306	3.8	1863	9	US-09-734-672-6	Sequence 6, Appli
28	305	3.8	1158	10	US-09-815-242-12611	Sequence 12611, A
29	304	3.8	6281	10	US-09-815-242-12396	Sequence 12396, A
30	296.5	3.7	1285	10	US-09-982-091A-2	Sequence 2, Appli
31	292.5	3.6	350	9	US-10-101-487-58	Sequence 58, Appli
32	290.5	3.6	1597	9	US-09-832-292-35	Sequence 35, Appli
33	285.5	3.5	2437	10	US-09-815-243-5834	Sequence 5834, Ap
34	284	3.5	1501	10	US-09-924-154-17	Sequence 17, Appli
35	281.5	3.5	1085	10	US-09-864-761-43248	Sequence 43248, A
36	279.5	3.5	2828	10	US-09-905-123-21	Sequence 21, Appli
37	279.5	3.5	2828	10	US-09-991-630-21	Sequence 21, Appli
38	278	3.4	2076	10	US-09-815-242-5815	Sequence 5815, Ap
39	278	3.4	2186	10	US-09-815-243-12913	Sequence 12913, A
40	277.5	3.4	2025	10	US-09-815-243-5703	Sequence 5703, Ap
41	274	3.4	858	9	US-09-684-001-19	Sequence 19, Appli
42	267.5	3.3	1332	10	US-09-982-091A-4	Sequence 4, Appli
43	266.5	3.3	1192	10	US-09-758-140-6	Sequence 6, Appli
44	266.5	3.3	1192	10	US-09-972-599A-6	Sequence 6, Appli
45	265.5	3.3	1192	10	US-09-789-386-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

```

US-09-902-432-4
; Sequence 4, Application US/09902432
; Patent No. US20020160002A1
; GENERAL INFORMATION:
; APPLICANT: Irwin H. Gelman
; APPLICANT: Susan G. Jaken
; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
; FILE REFERENCE: A30558-A-PWC-A 070156, 0597
; CURRENT APPLICATION NUMBER: US/09/902,432
; CURRENT FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 08/978,277
; PRIOR FILING DATE: 1997-11-25
; PRIOR APPLICATION NUMBER: 08/665,401
; PRIOR FILING DATE: 1996-06-18
; PRIOR APPLICATION NUMBER: 08/635,121
; PRIOR FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1596
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-902-432-4

```

[illegible]

|||||
Db 181 LTVKKDSEEGEABASVAGDHOBPVETAVGESASKESELKOSTEQESTLQOQOSTIIP 240
Qy 241 LQASDQAABEABADGEKEKEPTKSPESPSPVNSSETTSFFKFFTHGAGRRKTS 300
Db 241 LQASDQAABEABADGEKEKEPTKSPESPSPVNSSETTSFFKFFTHGAGRRKTS 300
Qy 301 FKSKEDDLLETAEKKEBAEKVDEEKEKTEPASEBOPADDTQALSLADYKVELPL 360
Db 301 FKSKEDDLLETAEKKEBAEKVDEEKEKTEPASEBOPADDTQALSLADYKVELPL 360
Qy 361 EDQVGDLEASSSEKCAPLATEVFEDEKMEAHQEVVAEVHYSTYEKTEEBOGGGGEAGGV 420
Db 361 EDQVGDLEASSSEKCAPLATEVFEDEKMEAHQEVVAEVHYSTYEKTEEBOGGGGEAGGV 420
Qy 421 VEGTGESLPEKTLAEPQVPOEABPAEELMKSRMCMVSGGDHTQTLTDLSPBEKTLPKAPE 480
Db 421 VEGTGESLPEKTLAEPQVPOEABPAEELMKSRMCMVSGGDHTQTLTDLSPBEKTLPKAPE 480
Qy 481 GIVSEVEMLSQOERIKVOGSPKLKLPSSSGLKLSGKKQKGRGGGDEBPEYOHYTE 540
Db 481 GIVSEVEMLSQOERIKVOGSPKLKLPSSSGLKLSGKKQKGRGGGDEBPEYOHYTE 540
Qy 541 SPESADBEKGESSASSPEPEETTCLEKGPLAPQDGEABEGTSDGEKKGAGITPMASF 600
Db 541 SPESADBEKGESSASSPEPEETTCLEKGPLAPQDGEABEGTSDGEKKGAGITPMASF 600
Qy 601 KKMVTPKRRVRPSESDDELEKYSATLSSTDSVSEMDEVTVGEBQKPEEPKRV 660
Db 601 KKMVTPKRRVRPSESDDELEKYSATLSSTDSVSEMDEVTVGEBQKPEEPKRV 660
Qy 661 DTSVSWEALICVSSKKCARAKASSDDDEGPTTLCGDHRAEASKDXKAGTDAVPASTQ 720
Db 661 DTSVSWEALICVSSKKCARAKASSDDDEGPTTLCGDHRAEASKDXKAGTDAVPASTQ 720
Qy 721 EDDQOQSSSPBPASSPSEBEGSVTWESFKRLVTPRKKSLSLEKAEADSSVEQLSTEIE 780
Db 721 EDDQOQSSSPBPASSPSEBEGSVTWESFKRLVTPRKKSLSLEKAEADSSVEQLSTEIE 780
Qy 781 PSRESWMSIKKFIPIGRKKRADGQEQATVEDSGPVEINEDPNVPAVPLSEYNAVER 840
Db 781 PSRESWMSIKKFIPIGRKKRADGQEQATVEDSGPVEINEDPNVPAVPLSEYNAVER 840
Qy 841 EKMEAQNTLEPOLLIGAVVSELSKTLVHTVSVAVIDSTRAVTSYEBRSPEWISASYTE 900
Db 841 EKMEAQNTLEPOLLIGAVVSELSKTLVHTVSVAVIDSTRAVTSYEBRSPEWISASYTE 900
Qy 901 PLEHTAGEAMPVEEVEVTEKDIIAETPVLTQTLPEGKAHDMVTSEVDFTEAVTALET 960
Db 901 PLEHTAGEAMPVEEVEVTEKDIIAETPVLTQTLPEGKAHDMVTSEVDFTEAVTALET 960
Qy 961 SPALTEETEASGABETTDWMSAVSOLTDSPTEEATPVQEVESGVLDTEEBERQTA 1020
Db 961 SPALTEETEASGABETTDWMSAVSOLTDSPTEEATPVQEVESGVLDTEEBERQTA 1020
Qy 1021 ILQAVADKVEESQVPAQTQVORTGSKALEKEVEEEDSEVLASEKXQVMPKGPVQENAG 1080
Db 1021 ILQAVADKVEESQVPAQTQVORTGSKALEKEVEEEDSEVLASEKXQVMPKGPVQENAG 1080
Qy 1081 AEHLAQSGFTGATPESLEPEVPTADVHVATCOVTKLOQLEQAVAPSSSETLTDSEIN 1140
Db 1081 AEHLAQSGFTGATPESLEPEVPTADVHVATCOVTKLOQLEQAVAPSSSETLTDSEIN 1140
Qy 1141 GSTPLADSDTAAGTOODETIDSQDSKATAAVASQVTEEBEATAQKEBESSTLPNNVPAOE 1200
Db 1141 GSTPLADSDTAAGTOODETIDSQDSKATAAVASQVTEEBEATAQKEBESSTLPNNVPAOE 1200
Qy 1201 EHGEERGRVLEPTQOELTAAAVPVIAKTEVQEGEVDMLDEKYEKEOEVFVHSGPNSQ 1260
Db 1201 EHGEERGRVLEPTQOELTAAAVPVIAKTEVQEGEVDMLDEKYEKEOEVFVHSGPNSQ 1260
Qy 1261 KAADVTYDSEWGVAGCQKEKSTEVQSLSEBEMETDVEKEKETKPEQVSEBGEQETA 1320
|||||

Db 1261 KAADVTYDSEWGVAGCQKEKSTEVQSLSEBEMETDVEKEKETKPEQVSEBGEQETA 1320
Qy 1321 APEHEGTGKPVLTLDMPSSERKALGSLGSPRLPPODKAGCTIEVVOQSIDTTVTQTAE 1380
Db 1321 APEHEGTGKPVLTLDMPSSERKALGSLGSPRLPPODKAGCTIEVVOQSIDTTVTQTAE 1380
Qy 1381 AVEKVIETVWISSETGESPECVGAHLPAEKSSATGHWTLQHAEDTVPLGPESQAESIPI 1440
Db 1381 AVEKVIETVWISSETGESPECVGAHLPAEKSSATGHWTLQHAEDTVPLGPESQAESIPI 1440
Qy 1441 IVPAPSTLHPDLQGEISASQERSESEEDKPDAPDADGKSTAIKVLKABPELLELE 1500
Db 1441 IVPAPSTLHPDLQGEISASQERSESEEDKPDAPDADGKSTAIKVLKABPELLELE 1500
Qy 1501 SKSNKIYLVNIQTAVDOPARTETAPETHAIDSQTVACRLDSEBPRCWTMADAKMKA 1560
Db 1501 SKSNKIYLVNIQTAVDOPARTETAPETHAIDSQTVACRLDSEBPRCWTMADAKMKA 1560
Qy 1561 PVPQREDLOVLTVLEAMAOPRKCLPRLQKAPVSK 1596
Db 1561 PVPQREDLOVLTVLEAMAOPRKCLPRLQKAPVSK 1596

RESULT 2
US-09-902-432-2
; Sequence 2, Application US/09902432
; Patent No. US20020160002A1
; GENERAL INFORMATION:
; APPLICANT: Susan G. Jaken
; APPLICANT: Irwin H. Gelman
; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
; FILE REFERENCE: A30558-A-FWC-A 070156.0597
; CURRENT APPLICATION NUMBER: US/09/902.432
; CURRENT FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 08/978,277
; PRIOR FILING DATE: 1997-11-25
; PRIOR APPLICATION NUMBER: 08/665,401
; PRIOR FILING DATE: 1996-06-18
; PRIOR APPLICATION NUMBER: 08/635,121
; PRIOR FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1346
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-902-432-2

Query Match 67.2%; Score 5421.5; DB 9; Length 1346;
Best Local Similarity 91.6%; Pred. No. 6.4e-229;
Matches 1098; Conservative 24; Mismatches 62; Indels 15; Gaps 7;

Qy 387 MEAHQEVVAEVAHVSTVEKTEEBEBOGGGGEABEGVVEGTGSLPEKTLAEPQVQEAPEA 446
Db 1 MEAHQEVVAEVAHVSTVEKTEEBEBOGGGGEABEGVVEGTGSLPEKTLAEPQVQEAPEA 60
Qy 447 EELMKSRMCMVSGGDHTQTLTDLSPBEKTLPKHPRGIVSEVEMLSQOERIKVOGSPKLKLF 506
Db 61 EELMKSRMCMVSGGDHTQTLTDLSPBEKTLPKHPRGIVSEVEMLSQOERIKVOGSPKLKLF 120
Qy 507 SSSGKTKLSGKKQKGRGGGDEBPEYOHHTSPESADQKESASSPEPEETTC 566
Db 121 SSSGKTKLSGKKQKGRGGGDEBPEYOHHTSPESADQKESASSPEPEETTC 180
Qy 567 EKGPLEAPQDEAELEGTTSDGEKKRE---GITPMASFKKMVTPEKRRVRPSESDKEBEL 622
Db 181 EKGPLEAPQDEAELEGTTSDGEKKRE---GITPMASFKKMVTPEKRRVRPSESDKEBEL 234
Qy 623 EKVSATLSTSDTSVSEMDEVKTVEBQKPEPRRVDTSVSMALICVSSSKKARAKA 682
Db 235 EKVSATLSTSDTSVSEMDEVKTVEBQKPEPRRVDTSVSMALICVSSSKKARAKA 294
Qy 683 SSSDDEGPTTLCGDHRAEASKDXKAGTDAVPASTQEDQOQSSSPBPAGSPSEBEG 742
|||||

```
Db 295 SSSDIR-GPRTLGGGQSRGQQRSDTAVPASTQEQQAQSSSPAGSPSGEG 353
Qy 743 VSTWESFKRLVTPRKKSLSLEEK-ABDSVEQQLSTIEPSSREBSWS--IKKPIPGRRK 799
Db 354 VSTWESFKRLVTPRKKSLSLEEKAGRTLVVAGGCPRLSRVRKEKNLGFPLRNSPPDGR 413
Qy 800 KRADGKQEOATVEDSGVEINEDDPNVPVAVPLSEYNAREKMEAGNTLPELLGAVY 859
Db 414 KGQMGQEOATVEDSGVEINEDDPNVPVAVPLSEYNAREKMEAGNTLPELLGAVY 472
Qy 860 VSELSKTLVHTVSVAVIDGTTRAVTSVEERSPSWISASVTEPLEHTAGEAMPVVEVTEK 919
Db 473 VSELSKTLVHTVSVAVIDGTTRAVTSVEERSPSWISASVTEPLEHTAGEAMPVVEVTEK 532
Qy 920 DIIAETPVLITQTLPEKGADHDMVTSEVDFTSSEAVTATSETSEALRTEEVTEASGAEETT 979
Db 533 DIIAETPVLITQTLPEKGADHDMVTSEVDFTSSEAVTATSETSEALRTEEVTEASGAEETT 592
Qy 980 DMVSAYSQLTDSPTTTEATPVQVRESGLVDTSEERQTAILOAVADKVKESQVPATQ 1039
Db 593 DMVSAYSQLTDSPTTTEATPVQVRESGLVDTSEERQTAILOAVADKVKESQVPATQ 652
Qy 1040 TVORTGSKALEKVEEVEDSEVLASEKEDVMPKGPVQEAAGAEHLAOGSETGOATPRESLE 1099
Db 712 TVORTGSKALEKVEEVEDSEVLASEKEDVMPKGPVQEAAGAEHLAOGSETGOATPRESLE 712
Qy 1100 VPEVTADVHVATCOVKIKLOQMEQAVAPESSETLTLDSETNGSTPLADSDTADTQOQDET 1159
Db 713 VPEVTADVHVATCOVKIKLOQMEQAVAPESSETLTLDSETNGSTPLADSDTADTQOQDET 772
Qy 1160 IDSQDSKATAAVROSQVTESEEAATAQKEEPSSTLPNNVPAQEEHGEERGRDVLPTQOQELT 1219
Db 773 IDSQDSKATAAVROSQVTESEEAATAQKEEPSSTLPNNVPAQEEHGEERGRDVLPTQOQELA 832
Qy 1220 AAAPVPLAKTEVGOEGEVDMLDGEKVKERQEVFVHSGPNSQKAADVTDSEVMVAGCQOE 1279
Db 833 AAAPVPMQKTEVGOEGEVDMLDGEKVKERQEVFVHSGPNSQKAADVTDSEVMVAGCQOE 892
Qy 1280 KESTEVQSLSEEGEMETDVEKEREKTKPEQVSEGEQETAAPHEGTYGKPVLTLDMPSS 1339
Db 893 KESTEVQSLSEEGEMETDVEKEREKTKPEQVSEGEQETAAPHEHRYGKPVLTLDMPSS 952
Qy 1340 SERGKALGSLGGSPSLPDQKAGCIEVQVQSLDVTVTQTAEAVEKVITVVISETGSPE 1399
Db 953 SERGKALGSLGGSPSLPDQKAGCIEVQVQSLDVTVTQTAEAVEKVITVVISETGSPE 1012
Qy 1400 CVGAHLPLPAEKSSATGCHWTLOHAEDTVPLGPESQAESIPITVTPAPESTLHPDLOEIS 1459
Db 1013 CVGAHLPLPAEKSSATGCHWTLOHAEDTVPLGPESQAESIPITVTPAPESTLHPDLOEIS 1072
Qy 1460 ASQERSEEDKPDAGPDADGCKESTAJEKLKAPPEILESKNKLVLNVIOATVDQFA 1519
Db 1073 ASQERSEEDKPDAGPDADGCKESTAJDKVLKAPPEILESKNKLVLNVIOATVDQFA 1132
Qy 1520 RTETAPTHAYDSQTPVACRLDSREPNRCWTMKMDAKMHPVQPPREDIQLVTLVLEAW 1578
Db 1133 RTETAPTHAYDSQTPVAMRLDSREPNRCWTMKMVKMKHPVQPPREDIQLVTLVLEAW 1191

RESULT 3
US-09-738-877-3
; Sequence 3, Application US/09738877
; Patent No. US20020015970A1
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan
; APPLICANT: Weiss, Stephen J.
; APPLICANT: Glynné, Richard
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF ANGIOGENESIS, COMPOSITIONS, AND METHODS OF SCREENING FOR ANGIOGENESIS MODULATORS
; TITLE OF INVENTION: SCREENING FOR ANGIOGENESIS MODULATORS
; FILE REFERENCE: A-69806/DJB/JJD
; CURRENT APPLICATION NUMBER: US/09/738,877
```

```
; CURRENT FILING DATE: 2000-12-15
; PRIOR APPLICATION NUMBER: US 09/637,977
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: US 60/148,425
; PRIOR FILING DATE: 1999-08-11
; PRIOR APPLICATION NUMBER: PCT/US 00/22061
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 3
; LENGTH: 1781
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-738-877-3

Query Match 53.2%; Score 4294; DB 10; Length 1781;
Best Local Similarity 57.9%; Pred. No. 1e-179;
Matches 959; Conservative 187; Mismatches 401; Indels 108; Gaps 35;

Qy 1 MGAGSSTEQRSRSPQ-PAGSDTPSELVLSHGHPAAEAS-GAAGDPADA--DPATKLPQKNG 56
Db 1 MGAGSSTEQRSRSPQ-PAGSDTPSELVLSHGHPAAEAS-GAAGDPADA--DPATKLPQKNG 60
Qy 57 QLSNVNGVAEQGDVHVQENQEQ-----EEVVDEVDQRESEDVREKDR 102
Db 61 QLSNVNGVAEQGDVHVQENQEQ-----EEVVDEVDQRESEDVREKDR 120
Qy 103 VEMAAANSTAVEDITKDGQOETSEITQIPASNNVVEVMQPAESQANDVGFKKVPKFGV 162
Db 121 DKEMATKSAVVDITDDGQENRN-IEQIPSSSENLEELTQPTESQANDIGFKKVPKFGV 179
Qy 163 FKFTVKDKNKEKSDTVQLLTVKKDEGEAEASVAGADHQPSPSVETAVGESAKSESELKQS 222
Db 180 FKFTVKDKNKEKSDTVQLLTVKKDEGEAEASVAGADHQPSPSVETAVGESAKSESELKQS 234
Qy 223 TEKQEGTLKQESSTIPIQAESDQAAEBAEKDEGEKKEKTKPESPSSPNSSETTS 282
Db 235 TEKPEETLKREQSHABISPPAESQAV-EECKEKEGKQEKESKSAESPTSPVTSETGS 293
Qy 283 SPKKFTTHGWAGWRKTSFKKSKEDDLETAERKEQEAQKVDDEEKEKTEPASEE----- 337
Db 294 TFKKFTTQGWAGWRKTSFKKSKEDDLETAERKEQEAQKVDDEEKEKTEPASEE----- 353
Qy 338 -----QBPADTDQARLSADYKVELPLEDQVGDLEASSEEKCAPLATEVFEDERMEAHQ-E 392
Db 354 QAHQPQPAESAHEPRLSAEYKVELPSEEQVSGSQSQSEKPAPELATEVFEDKIEVHQEE 413
Qy 393 VVAEVHVSTVEKTEEREGGGEAGGVVVEGTGESLPPEKLAEPQVQPAEABELMKS 452
Db 414 VVAEVHVSTVEERTERQ-----KTEVEETAGSVPAEELVGMADAEQPAEPAKLVKL 465
Qy 453 REMCVSGGDHTQLTDLSPPEKTLPHKEGIVSEVEMLSQERIKVQGSPLKLFSSSGLK 512
Db 466 KETCVSGEDPTQOQADLSPEKVLKSPGEGVSEVEMLSQERIKVQGSPLKLFSSSGLK 525
Qy 513 KLSGKKQKRGKGDDGDEEPGEYQHIHTESPESADEQKGESSASSPEPEBTTCLEKGPLE 572
Db 526 KLSGKKQKGR-KGGDEESGEHTQVPADSPDSQEEQKGESSASSPEPEBTTCLEKGLAE 584
Qy 573 APQDGAEBEGTTSDBGKKREGITPWAISFKKMTVPKRVRRRPSSEDKKEELEKVKSATLSS 632
Db 585 VQDQGEAEAGATSDGKKREGITPWAISFKKMTVPKRVRRRPSSEDKKEELEKVKSATLSS 644
Qy 633 TDSVTSEMDEVKTVGEEOKPEEPKRVDTSVSWEALICVGSKKKARKASSSSDDEGGPR 692
Db 645 TESTASEMQEEMKGSVEEPKPKKVDTSVSWEALICVGSKKKARKASSSSDDEGGPR 704
Qy 693 TLGGDSHRAEASKDEAGTDAVPASTQEQDQAGQSSSPGAGSPGEGVSTWESFKRL 752
Db 705 AMGDHQKADACKDETGTGDLGAGQEHDPQCGSSSPGAGSPGEGVSTWESFKRL 764
Qy 753 VTPRKKSLSLEKAEAD-----SSVEQLSTEIPSRRESWSVIKKFTIPGRKKKADGKQEQ 808
Db 753 VTPRKKSLSLEKAEAD-----SSVEQLSTEIPSRRESWSVIKKFTIPGRKKKADGKQEQ 808
```



```

Db 765 VTPRKSKSLEEKSESDSIAAGSVHSTPDPTEBGEBSWVSIKFKLPGRRKKRPDQKQ 824
Qy 809 ATVEDSGVEINEDDPNNPVPVPLSEYNAAVEREKME--AQGNTLPLQLGAVYSEELS 865
Db 825 APVEDAGPFGANEDSDVPAVVPVLESDYAVEREKMEAOQAQGAQPOEOKATEVSKELS 884
Qy 866 KTLVHTVSAVAVDGRRAVSVERSPSMISASVTEPLETAGEAMPVPEVEKIDIA-E 924
Db 885 BSQVMMMAAAVADGTRAAITIEERSPSMISASVTEPLEVEKBAALLTVEVERETIAE 944
Qy 925 ETPVLTQTLPECKDAHDMDVTSVDFTSEAVTATETSEALRTEEVTEASGAETTDMSA 984
Db 945 BEPTVTEPLEPKREARAGDTVVESEALTEPAVTAETAGPLGSEEGTEAAGAEETEMVSA 1004
Qy 965 VSQLTDSPTTEBATPVQVESGVLDTEEEERQTAIIOAVADKYKEEVSQVAT----- 1038
Db 1005 VSQLTDSPTTEBATPVQVEGVDPDIEQERRTOEVLQAVAEKVEESQLTGTGPEDV 1064
Qy 1039 -QTVQRTSKALEKVEEVESEVLASEKEKDVMPGVPQVQGAHEHLAQGSFTGQATPS 1097
Db 1065 LQPVQRA-----EABRPEQAEASGLKKEVTVLKVDAQEKTEPFTQGVVGGTTPES 1118
Qy 1098 LE-VFETPADVDH--VATCOV-----IKLOOL-MEOAVAPESSELTLDSETNGSTPLAD 1147
Db 1119 EKAKQVETSISSSELYTTQCAETLAGVKSQEMMEQALPDSVETPTSTDTGSTRPAD 1178
Qy 1148 SDPADGTQODETIDSDSKATAVARSQVTEEAATAQKEEBSLTLPNNVPAOEHEGEPG 1207
Db 1179 PDPGTQODEIVEIHEENEVASGTQSGTEAABVAPAKERPRA--PSSFVQOEETKEQSK 1237
Qy 1208 -RDVLEPTQOELTAAAVPLATTEVQGEVDMQEKY-----BEQVFVHSG--PMS 1259
Db 1238 MEDTLEHTDKSVSVETVSLSTKTEGTQ--EADQVADKTKVDPFFEGLEGSIDTGIYVSR 1295
Qy 1260 OKAADVTYSEVWVAGVCOEKSESTEVOS-----LSLEEGEMTEVDEKEREKTEPEQVSE 1314
Db 1296 EKVTEVALKGEETEEBCKKDALBELQSHAKSPSPVEHEMVQVREKTEAPHTVNEE 1355
Qy 1315 G-EQETAAEHEGTYGKPVLTLDMPSSERKALGSLGSS--PSLPDQKACIEVOYQSID 1372
Db 1356 KLEHEETAUVSESVSKQLQTVNVPPIIDAKESVLESGPPCLQGEFAVCTKIQVQSSSE 1415
Qy 1373 TTVTQTAAEVKEVI--ETVVISSETGSPCVGAHLLPAKKSATGTHMLQHAEDVPIG 1430
Db 1416 AFTLTAAAEKVKVGETVNILETGTELPAGAHVLEKSEKONEDFAHPGEEDVPTG 1475
Qy 1431 PSSQASIRIITYTPAPESLTLPLOGEISASQERSEEDKPDAGPDAGKESTALEKVL 1490
Db 1476 PDCQAKSTVIYSAITTKKLSDBEGEKTTSLKMSDEVDEVOVACQEV--KYSVAILED- 1532
Qy 1491 KAEPE--ILEESKSNKIYLVNIQTAVDQPART-ETAPETHAVDSQTVQVPAICRLSREPN 1547
Db 1533 --EPENGILELETKSKLYONIIQTAVDQFVTRTEETATEMLTSELTQAHVIAKADQDAG 1590
Qy 1548 KCMWK-----MDAKMKHPVQPREDLQVLYLEA 1577
Db 1591 QETKEGEPPOASADEPTPITSAKESSESTAVGOA 1625

```

```

RESULT 4
US-09-742-096-3
; Sequence 3, Application US/09742096
; Patent No. US20020155441A1
; GENERAL INFORMATION:
; APPLICANT: DRUIHE, PIERRE
; APPLICANT: DAUBERSTES, PIERRE
; TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
; FILE REFERENCE: 200773USODIV
; CURRENT APPLICATION NUMBER: US/09/742,096
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 08/973,642
; PRIOR FILING DATE: 1998-02-06
; PRIOR APPLICATION NUMBER: PCT/FR96/00894

```

```

; PRIOR FILING DATE: 1996-06-12
; PRIOR APPLICATION NUMBER: FR 95/07007
; PRIOR FILING DATE: 1995-06-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1786
; TYPE: PRT
; ORGANISM: P. falciparum
US-09-742-096-3

Query Match 5.7%; Score 462; DB 9; Length 1786;
Best Local Similarity 19.2%; Pred. No. 5.6e-13;
Matches 325; Conservative 291; Mismatches 644; Indels 436; Gaps 69;

Qy 53 QNQGOLSVNGVABQGDVHVOENQGOE-EVYDEVDQGRESDEVREK--DRVE-- 105
Db 95 QVNGELAS-----EVKEKIIDLLEGGNTLTESYDDNGLNEADIKENILSNIEPKE 149
Qy 106 -----MAANGTAVEDITKDGQ---EETSEIIEQIPASENNVEENVQPAEQANDVG 153
Db 150 NIIDLNNIGNQSEKQESVENQVSDLEFNEELNSVDVNGEYKENTLE--ESQVNDI 207
Qy 154 FKVPKFFGFKFTYKDKNEKSDTVQLLTVKKDGEAGASV--GAGDHPSTVETAYGE 211
Db 208 FNSLYK-----SVQEQQHVAEKEVESVEENDESVAEENVEENDESGVASSVEE 261
Qy 212 S--ASKESSELKOSTEKQGTIKQKQ-----QSTETPLQAESDQAAE 251
Db 262 SIASVDSIDSIIEENVAFTVEIIVAPSVESVAPSVESVEENVEESVAENVEESVAE 321
Qy 262 EAKDGEKQKEPEKPSPPSSPNSETTSSFKKFTFHGAQRKKTSPKSKSEDDLE- 310
Db 322 NVESVVAENVEESVAENVEEIVAPFTVEIIVAPSVESVAPSVESVEENVEE 381
Qy 311 -TAEKREKQEAKEYDEEKEKTEPA---SEQOEPADTDQARLADYKVELPLEDQVGD 366
Db 362 SVAEENVEESVAENVEESVAENVEESVAENVEESVAENVEEL-VAPFTVEIIVAPFTVEI- 438
Qy 367 LEASSEKCAPLATEVFDEKMEHQEVAEVAHVSTVEKTEEEOGGGGRABGGVVEGTGE 426
Db 439 VAPSVESVAPSVESVEENVE--ESVAENVEESVAENVEES-----VAENVEE 485
Qy 427 SLPPKLAEPQEVPOEAPPAEELMKSRMCVSGGDHDTLTLSPBEKTLPHHPBEIVSEV 486
Db 486 SV-----AEVVEESVAE-----NEESVAENVEESVAENV 515
Qy 487 EMLSSQERIKVQGSPLKPLFSSGLKLGKKQKGRGGGDEBEGEYQIHITESPESAD 546
Db 516 BEIIVAPFTVEIIVAPFTVEIIVAPSVESVA-----PSVESVEENVEESVA 560
Qy 547 EQKGESSASSPEEP-----EETTLCEKGPLAPQDGAEBGTTSDGKKREKIGTPWMAFK 601
Db 561 ENVEESVAENVEESVAENVEESVAENVEEIVAPFTVEIIVAPFTV-----EIVAP--SVV 612
Qy 602 KMVTKKAVRRPSESDKEEL-EKYKSAITLSSTDSVTSEMODEV--KTVGGEQKPEEKR 658
Db 613 BSVAPS--VEESVEENVEESVAENVEESVAENVEESVAENVEEIVAPFTVEIIVAP- 665
Qy 659 RYDTSVSMELICVQSSKKRRKAKSSSDDEGGPRLGDSHRAEF--ASKQKEKGTADVP 716
Db 666 -----TVEEIVAPSVESVAPSVESVEE-----NVEESVAENVEESVAENVEE 708
Qy 717 ASTQEQDQAGSSSPPEPASPSEGEVSTWESFKLVT-----RKKSKEKLEKADS 770
Db 709 BSVANVEESVAENVEEIVAPFTVEIIVAPFTVEIIVAPSVESVAPSVESVEENVEES 766
Qy 771 SVEQLSTIEPSEESQVSIKKFTPGRRKKRQADGQKQQA--TVEDSGVEINEDDPNVA 828
Db 767 VAENVEESVAENVEES-----VAENVEESVAPFTVEIIVAPSVEE-----S 806
Qy 829 VVPLSEYNAVEREKMEAGNTLPLQLGAVVSEELSTLVHTVSVAVIDGRATVSVEE 888

```



```

QY 168 KKDNEKSDTVQVLTKVKBGBGAEASVAGDHPESVETAVGBASAKESSELOKSTOKO 227
Db 436 -----KLDRLNTVASPXDQCELASISVSGSGRPSSDLOARLIGELAG-----ESVENOE 483
QY 228 GTLKQOOSTELP-----LOAESDOAAEEB--AKDGBEEKQEKPEPTKSPSPSPNS 278
Db 484 -----VQSKKPLPSKRPOLKQLOVLDQGBEREDVRKNYCSLDELTPERKSGQEKSHSVNT 538
QY 279 ET-----TSSFKKFTTHGAWGRKKTSPFKSK-----EDDLETAEKR----- 315
Db 539 EFKIGIDIDHTQSVKQWQEQRRKQOQMEELAKSEKFGSPKQDVVEYERRSLVHEVGKP 598
QY 316 -----KEQBAEKVD-----EBEKEETPASEBOEPAEDTDOARLADYEKVEL 358
Db 599 QDVTDSPPSKKRMHDVDFDICTRERNYRSSQISBESERTGSPSVRHGSHFED-- 655
QY 359 PLEDOVGDLAESS-----BEKCAP-----LATEVFDEKMEAHQEVAVHVS 401
Db 656 --EDPIGSPRLISVKGSPKVDKVLPSYNTTYRESLKNPFDSSRRQOMADMAKIKLSV 713
QY 402 VEKTEEOGGGGGAE--GGVVEGTGESLPPEKL-----AEQEVPOEAPAE-- 448
Db 714 LNSEBELNRWDSQMDQADRFVSPNSIIRKDSLKRSVRLDEPGEVPSDDEDEGHKS 773
QY 449 -----LMKSRMCVSGGDHTQLTLSPE-----EKLPHKPGIYSE 485
Db 774 HSPRASALYESRLSFLRLRREDKLERDERLSSSLERNKFPYFALDKITPTDTALLER 833
QY 466 VEMLSS-----QERIVQGP-----LKKLFSSGGLKXLSG 516
Db 834 AKSLSSSREENWSFLDWDSRFANFRNNKDKEXVDSAPRIPSWYMKKKIKRTDSSGKMD 893
QY 517 KQOKKRGGGGDEEGEYOHITESPESADEQKGSASSPPEPTTCLKGPLEAPOD 576
Db 894 KKEDHK-----EEBEOROLFASRFLHSSIFEODSKRLQHLRKEBDSPIGRLYGQT 948
QY 577 GBAEGGTSDEGKKEGITTWAS--FKWV-----TPKGRVRPSESDEKELEKVKASIS 631
Db 949 SEGANSTT--DSIQEPVLFHSRFMELTRMOQKEKQKKEVEKQEDTNHPTPERS 1005
QY 632 STDSIVSEWQ-----DEVKTVEGQKPEPKRVDTSVSEWALICVSSSKKARAKAS 684
Db 1006 AENKDSLEKTPPSVGPSTVVTLESAPSALEKTTGKTVEAPLV--TEETVVEPAPV 1062
QY 685 SPD-----EGCPRTLGGDSHRAE-----ASKDKEAGTDAVAPASTODQOAGSSPBP 733
Db 1063 SBEAKRASPAPAPV-----EQLEQVLDLPPGADPDKEAA--MMPAGVE-----GSSGDQP 1111
QY 734 ---AGSPSEGEVSTWE-----SPKRLVTPRKKSQKLEKAE--DSVVE----- 773
Db 1112 PYLDKAPPTPGASFQAEENVDPEPDSTPLSKPAQKSEANEPRKAEKPDATADAPDAN 1171
QY 774 ---QSTELPESREBSWISIKKFIPIGRKKRADGQEOQTVEDSGVEINEDDPNPAV 830
Db 1172 OKAEAPESQPPASED--LEVDPVPAKDKPKNSKSKTPVQ--AAAVSVI--KPVT 1224
QY 831 PSEYNAVEREKMEAOQNT--ELPOLLGAVVYSEELSKT-----LVHTVSAVADGT 880
Db 1225 RKSF--RIRREKLRKNSRGEQKLELKMEXKITRTASKNSAADLHPESLPLSKT 1282
QY 881 RA--VTSV-----EERSPSWISASVTEPLEHTAGEAMPVVEVEKQILAE--ETPV 929
Db 1283 RRRNRVSVAATMGDHNRSR-----VKEPVEQ-----PRVTRKLEBELQAAAV 1327
QY 930 TGT-----LPEGDAHDWVTSEVDTSEAVNTATSEALRTBEVEEAAGAE--ETTD 980
Db 1328 PTPPRGRPPKTRRRARDEEENAKAPAEFTLKPEGMRSRSPQKTAAGGPOQKKNBP 1387
QY 981 MYSASQULDSPTTEAPAPVQGVESGVLDTEEBEROTQIILQAVADKVYESQVATQT 1040
Db 1388 KVDAT-----REPAITTEVGPQIGVKESSMPEKAAEBEAGEQKRRDKAGTDCNPPETAP 1442

```

```

QY 1041 VORTSKALEKVEEVEDS-----EVLASEKED--VMPKGPVOEAGAEHL 1084
Db 1443 VE-----VEKKAPPEKNSKSGKGRSNSRLAYDKASLSKVNDAASPRGAAAGER-- 1495
QY 1085 AOGSETG--QATPESLVEPVTADVDHATQVYIKQOUMEQAVAPESSETLTDSETNGS 1142
Db 1496 -----ESGVAVVSPKESSPQKEDOLS-----SQLKSDPVDPKPEKEKEDVASAGP 1541
QY 1143 TPLADSTAGTQODETIDS-----QSKATAAARQS-----QYTEEAAT 1183
Db 1542 SPEA--TQIARQWELQAVEHIAIKLAESASAAAYKADAPBGLAEDRDKPAHQASETLAA 1600
QY 1184 AQK---EESTLPPNNVPAQEH--GE-----PGRDVLPTQOEL--TAAAVPLAKTEV 1231
Db 1601 AIGSTINDISEPNEFPAPPPYEPESQTDQOPRAGQALQPSSEGMTTDAVSSILETEA 1660
QY 1232 QOGEVMDLGEKYEQOEQEVFVHSGPNSQKADVTYDSEWVAGCQEKESTEVOSISLE 1291
Db 1661 ATESRPPVNAPODS-----AGPTDTKEARGNSSETSHSVPEAKSGKEVEVTLVRKD 1712
QY 1292 EGENETVEKREKRETKEQYSEBGEQETAPRHEGTYGKXVLTILDMSSRKGALGSLGG 1351
Db 1713 KGRQKTRSRKRNKTNKKVVA--PVESHVPSNOAG--ESPANEGTTV--Q 1759
QY 1352 SPSPDODKAGCIEVOYOSLDTVTYOTA-----EAVEKVIETVVISETGESPECVGAH 1404
Db 1760 HPEAPQEBK-----QSEKPHSTPQOSTDLSKIBSTENSSQGISVEBKPTPTASVPPD 1813
QY 1405 L-----LPAEKSSATGHWTLQHAEDTVPLGBESQAESIPILVTPA----- 1445
Db 1814 LPPPPQAPVDEEQAPRVHISIESDPVTPPSPD--PSIPPTLSVTAALKSPVVASGI 1872
QY 1446 PESTHPLQCEISASQRENSSEEDKPDAGDADGKE---STAIKVLAEPIILELES 1501
Db 1873 PHOSPPTKVTMTIROBEPRAQSTPSPALPBDTKASVDTSSSTRKITL--MDPYVATIS 1931
QY 1502 KSNKVLNVITQAVDQFARTTAPETAHAYBSQTPVPCARLDSREPNCWTMKAKMKHP 1561
Db 1932 VTSISVTTALAEV-----SAPCLH-----EAPPPVDSKAP-----LEEKTAAP 1972
QY 1562 VPOPREDLQVLTLEANAQPRKCLPRLQK-----APVS 1595
Db 1973 VTN--NSEIQASEVLVA--ADKEKAVPVIAPKITSVISMPPVS 2011

RESULT 6
US-09-815-242-5635
; Sequence 5635, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA 011A
; CURRENT APPLICATION NUMBER: US/09/815, 242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27

```

```
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5635
; LENGTH: 2368
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5635

Query Match          5.0%; Score 407.5; DB 10; Length 2368;
Best Local Similarity 17.5%; Pred.No.1.8e-10;
Matches 317; Conservative 293; Mismatches 663; Indels 541; Gaps

QY 47 PATKLPOKNGQLSVNGVAEQGDVHVQEENQGQEE-EVDVEDGQRESEDVRKD---101
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 102 RVEMAANSTAVEDITKDGQ-----EETSEI----IEQIPASENNVEEMVQPASQ 148
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 733 NAQVEAIKTKAINDINOTAPATTAKAAALEEFDEVVQAIDQAPLNPDPDTTNEEVAER 792
QY 149 AN--DVGFKKVKFVGFKFTVKDKNKSQDVTQLLT---VKDEGEAEASVAGQHQE 202
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 793 INAAKSVGVAIEATTAQDLERVKNKEISKIENITDSTQTMDAYNEVQOATARKTON 852
QY 203 PSVETA VGESAKSESLKQSTEK-----QEGTLQEOSSTE-----IPLQAESDQ 247
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 853 ATVSNATNEEVAADAAEAAGQGLHDLOVVKSQEVADTKSKVLKINKAIOQAQVKP 912
QY 248 AABEEAKDE-GEEKQKEPTKSPESPSPVNSETTSFFKPFTHGWAGWRKTSFKKSK- 305
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 913 AADTEVENAYNRKQEIQNS-----NASTTEEKQAAYT-----ELDTKKQEA 954
QY 306 -----EDDLETAB-----KRKEGBAEKVDEEBEKT-----EPAS 335
DB | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 955 RTNLDAANTNSDVTTKNDNGIAINOVAATTKKSDAKAEIAQAUSERKTAIEAMNDSTT 1014
QY 336 EEOPEAED-TQARLSADYEKVELPLEDOVDGLEASSEKCA-----PLATEVFDE 385
DB ||| : | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | |
QY 1015 EQQAQAKVDQAVTANKADINDNAANTDVDNAKTTNEATIAATPDANKVPATAQAIAD 1074
QY 386 KWEAHQEWAEVHVSTVEKTEEBQQGGGAEGGVVVEGTSGESLPPEKLAEPQEVQEAEP 445
DB | : | : | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | |
QY 1075 KVQAQETA I-----DANGATTE-----EKAANAQVOVTEKTT 1107
QY 446 ABELMKSRMCVCGGHDTQLDLSPEKTLPKHPGIVSEVMJSSQBRIKVGQSPLUKL 505
DB | : | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | |
QY 1108 ADTA I-----DGAHTN-----AEVEAAKNAEIAKIEAIQPAT 1140
QY 506 FSSGLKLKGKQKGRGGGDEEPGEYOIHTE-----S 541
DB | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 1141 TKONAKAIAITKANERKTA-----IAQTQDITASEIAAANAANAVDNNAVTOANNNIEANS 1194
QY 542 PESADEOK--GESSAS--SPEEPEETTCL--EKGPLEAPQDGEAEEGTTSDEKKRGKIT 595
DB | : | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | |
QY 1195 QNVDDVQAKTIGEASIDOVTFPVNKKATATVDKANNITAAATDNDGVDTAKOAGRNSIQSTQ 1254
QY 596 PWASFKK-----MVTPKRVRPRPSDSKEE----- 620
DB | : | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | |
QY 1255 PATAVKSNAKNDVDQAVTTQNQAIDNTTGATTEENAKADLVLKAKEKAYODILNAQTNN 1314
QY 621 ELEKVKSATLS-----STDSTUSEM-ODEVKTVGBEQK-----PEPKRRVD 661
DB | : | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | |
QY 1315 DVTQIKQADVQGITADTTIKDVAKDELATKAREOKALIAQTADATTEEKQANOQVD 1374
QY 662 TSVSW-----EALICVGSSKKRKARKA-----SSSDDEGGPRT-----LG 695
DB | : | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | |
QY 1375 AELTQGNQNTENAQSIDDVNTAKNDIAQIDPIQAQTDVKTNARAELLTEMQNKITILN 1434
QY 696 GDSHRAEASKD-----KEAGTDVAPAST-----OEQDAQSGSSPEPAG 735
```

RESULT 7
US-09-815-242-12389
; Sequence 12389, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12389
LENGTH: 2368
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-12389

Query Match 5.0%; Score 407.5; DB 10; Length 2368;

Best Local Similarity 17.5%; Pred. No. 1.8e-10; Indels 541; Gaps 71;
Matches 317; Conservative 293; Mismatches 663;

47 PATKLPKRGQSLSYNGVAEGDVAHQEENOGQEE-BVVDVVGQRESEVDREKD--- 101
Db 677 PTVKQAQKODIIQAVTTRKQ---IKSNASGLQDEKDVANRKGIEFKAIKIDAAAT 732
Qy 102 RVEEMANSTAVEDITKQO---EETSET---TEQIPASENNVEEMQPAESQ 148
Db 723 NAQVBAIKTAINDINQTPATTAKAALAEFEDEVQAQIDQAPLNPDTTNEVEAEIR 792
Qy 149 AN--DVGFKVKKFVGKFTVKKKDKREKSDVQULL---VKKDEGEAGAVGAGDHOE 202
Db 793 INAAKSGVKALEATTTAODLERVKNEISKIENTIDSTQTKMDAYNEVKAATARKTON 852
Qy 203 PSVETAVGESASELSKOSTEK---QEGTLKQEOSSTE---IPLQAESDQ 247
Db 853 ATVSNAKTNEVEAADAABAQAQGLHDQVVKSKQEVADTSKVLKIDNAIQTAQAKVP 912
Qy 248 AAEBAKDE-GEKQEKPTKSPESPSPVNSETTSPFKKFTTHGMAKRRKTSFKSK- 305
Db 913 AADTEVENAYNRKQEIQNS---NASTTEKQAAYT---ELDTKQKOA 954
Qy 306 ---EDLETFE---KKEQEAKEKDEBEKEKT---EPAS 335
Db 955 RTNLDAANTNSDVTAKGNGIAINOVQAATTKKSPAKAKELIAOKASERKTAIEAMNDSTT 1014
Qy 336 EEOQPAED-TDQARLSADYEKVELPLEDQVGLLEASSEKCA---PLATEVPDE 385
Db 1015 EEOQAAKQKVDQAVTANADINDMAANTVDNAKTTNETTIAITPDANVKTAKQALAD 1074
Qy 386 KMEAHQEVVAEYHVSTVEKTEEBEGGGGEAEQGVVEGTGESLPPKTLAPQEVPOEAP 445
Db 1075 KVOAQETAI---DANNQATTE---EKAAKAQOVQTEKTT 1107
Qy 446 AEELKSRMCVSGGDHDTQDLSPEEKLPHKPBQIVSEVEMLSQGEIKVQGSLLKTL 505

Db 1108 ADTAI-----DGAHTN-----AEVEAAKNAIEAKIAIOPATT 1140
Qy 506 FSSGKLKLSGKKQKGRGGGDEPGEYQHTE-----S 541
Db 1141 TKDAKAKAIATKAERKTA-----IAQODITAEERIAAANAVNADVQAANNIEANS 1194
Qy 542 PESADEOK--GESSAS--SPEPEETCL--EKGPLEAPQDGEAEQTSDEKKEKRGIT 595
Db 1195 QNDVDQAKTTGEASIDQVTPVVKKATAVDAKNNITAAIDNDGVDTAKDAGKNSIOSTQ 1254
Qy 596 PWASFKK-----WTPPKRVRRPSPEKDE----- 620
Db 1255 PATAVKSNAKNDVQAVTTQQAIDNTGATTEKNAKDLVLAKKXAVQDILNAQTNN 1314
Qy 621 ELEKKSATLS---STDSTVSEM--ODEYKTVGEBOK-----PEEPKRV 661
Db 1315 DVTOIKQAVADVQGITADDTIKVADDELATKAREKALIAQTPADATTEKEQANOQVD 1374
Qy 662 TSVSW-----BALICVGSKKRKARKA-----SSDDEGGPRT-----LG 695
Db 1375 AELTQGNQNIENAOISIDVNTAKDNALQALDPIQASTDVKTNAABELLTEMQNKITELN 1434
Qy 696 GDSHRAEASKD-----KEAGTDAVPAST-----QEDDOAGSSSPERAG 735
Db 1435 NNETTNEKNDIGPVRAAYEEGLNINAAITTDGVTAKOTANQVQQLHANVPKKPAG 1494
Qy 736 SPSEGEVSTWESFRRLVTPPKSKSKLEEKAE--DSSVEQLSTIEPSRESWV----- 788
Db 1495 KTAIDQAAAD--KKTQIEQTPNASQOEINDAKQOEVDTELNQAKTNIDOSTDEYDNAVKE 1553
Qy 789 -----SIKKFIPGR-----KKRAG 804
Db 1554 GKAKINAVKTFSEYKQALAKIEAIVNAKTEADNSNASTSEIAEAKQKALAEIKQIDNQ 1613
Qy 805 KOQATVEDSGPVEINEDPNV-----PAVPLSEYNAVEREKMEAOQNTLEPQ 853
Db 1614 NVNQATSKDIDLEVQIHNDLDINDYTIPTGKKEASITDLVAVADQKKNISADINATQDE 1673
Qy 854 LLGAVYSEELSKTLVHTVS-----VAVIDGTRAVTSVEERSPSMISASV 898
Db 1674 KQQAIIKQVQVQALAESINNGVNDGVDALTOGKAIDAIQVADATVPKPAQVVIDAKA 1733
Qy 899 TEPLP-----HTAGEAMPVEEYTE--KDIIAEFTV--LTQTLPGKXADHMV 944
Db 1734 EETKESIDQSQLTAETKTEKLAMIQTQDAKQGITDAITTAVERAKAGGLAFDVI- 1792
Qy 945 TSEVDTSEAVTATETSEALTEETVEASGAETTD-----MVASQQLTDS 992
Db 1793 --QIDSTKQKAIIELETAALQIEAGVNVADATTEBEKAFTNALLEDLSKATDISQT 1850
Qy 993 DTTEBATPVQEVESGLDTEEBERQTAI---LOAVADKVEESQV-----PATQTV 1041
Db 1851 TNAEIAIAT---VKNSALEQLAKRINPVPKKNALAEIAREVNNKQIEIIRKADADASAKEI 1906
Qy 1042 QRTG-----SKALEKVEEVEEDSEVLASEKEKDVMPK---GPQGEAGAEHLAQS- 1089
Db 1907 ARTDLGRYFDLPADLDLTQNTTEV--AELQNVTTIPALIEAIVPONDPNANDTNSGSDND 1964
Qy 1090 -----TGQATPESLEVPETADVVDVATQVYIKLQQLMEQAVABESSSETLTDSETNSTP 1144
Db 1965 ATANSNAAATPENGQPVHTESTDANAN-----DTSGTTNNQNDATG 2008
Qy 1145 LADSDTADGTQOQETID--SQDSKATAAVROSQVTEBAATAQKEEPTLPNN---VPA 1198
Db 2009 ETTATSANSATDANDKXPQANNNSADTSTNSPTMDVTSKEPEEST--NNGTTDXPA 2066
Qy 1199 QEHGEEBGRVLE-----PTQOELTAAAVAVLAKTBVGGQGVQDMDLGEKVK 1246
Db 2067 TEADNATPAESATNNSTTTTATNENAPTESTPATY--PTTAVST--GAESSADSKONASVN 2121
Qy 1247 E-EQEVFVHSG-----PNSQAKADVTVDSQVAVGAGQOEKESTVQSLSLEG 1293
Db 2122 DSKQNAEYNNASQSTNGKTAQPKSENKAKAEKDRGDRSTQSNVESTTTERLPADITEP 2181

Qy	1294	EMETDVEKEKRETKPQVSE--GEQTAAPEHEGTG-----KPVLTLDPSSER	1342
	:	: :	:
	:	: :	:
Db	2182	KVSNTSKDKKEESTTSOTDAEQHNSDTNVASNEADKSGNVDTDVSNKP--STSKPSEAK	2239
	:	: :	:
	:	: :	:
Qy	1343	GKALGSLGGSPSLPDODKAGCIEVQVSLDTVTTOTAAEVKIETVVVISSETGESPCVG	1402
	:	: :	:
	:	: :	:
Db	2240	DKA-----TSTDQSOKDMATADTKNQASIGATADVNNKATQN-----DG	2280
	:	: :	:
	:	: :	:
Qy	1403	AHLHPAEKSSATGGHTLQHAEYDTPGLPGES-QAESPIIIVTPAPESTLHPDLQGEIS	1461
	:	: :	:
	:	: :	:
Db	2281	ANASPA-----TVSKGSNSANQMILNVTKTENKANAKSAQQGVN--	2321
	:	: :	:
	:	: :	:
Qy	1462	QRERSEEDKPDAG	1475
	:	: :	:
	:	: :	:
Db	2322	-KPKQAKTLPDTG	2334

```

RESULT 8
US-09-864-761-35612
; Sequence 35612, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aomica-x-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 35612
; LENGTH: 1325
; TYPE: PRT

```

```
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AP000511.1
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.98
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.98
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EST_HUMAN HIT: AW867076.1, EVALUATE 1.00e-103
; OTHER INFORMATION: EST_HUMAN HIT: AW867076.1, EVALUATE 1.00e-103
; OTHER INFORMATION: SWISSPROT HIT: Q0Z817, EVALUATE 5.00e-63
US-09-864-761-35612

Query Match      4.8%; Score 388; DB 10; Length 1325;
Best Local Similarity 20.3%; Pred. No. 6.6e-10;
Matches 326; Conservative 191; Mismatches 662; Indels 428; Gaps 58;

Qy 53 QKQGQLSSVNGV-----AEQQDVHVQEEN-QEQBEEV 84
   |||::||: : : : : : : : : : : : : : : : : :
Db 8 QKTGELSTSDHIFPLTGLVYSIPFDHIVLHSGQRPELPKSTEIHEKGRCNTRHSKP 67

Qy 85 VDEDVGORESEDYREKDRVEEMAANSTAVEDIKDGOEETSEIIEOIPASENNVEEMVQP 144
   |||::||: : : : : : : : : : : : : : : : : :
Db 68 TDKRPTG--NSKITDHKSSTONHAPPTSSEENSQGGKD-----PMIRN--QRSVDP 114

Qy 145 AES---QANDVGFKVPKFVGFVFVKVDKNESDVTQLLLTVKKDEGEAASVGADHQ 201
   |||::||: : : : : : : : : : : : : : : : : :
Db 115 ADSTTHKEAGKHITPAKSKINCRRKSTTGKS----TVTRKSDK-----TORPL 161

Qy 202 EPSVETAVGSASKESELKQEGTLKQEOSSTEIPLQAESDQAAEBEAKDEGBEKQ 261
   |||::||: : : : : : : : : : : : : : : : : :
Db 162 EKSMSTL--DKTSTSCHKTTTSFHNSCNSQTOKSTSPF-----EKITAASKT 207

Qy 262 EKEPTKSPSPSPVNSETTSSPKFTTHGWAGRKTKTSFKKSKEDDLTAERKRKEQAE 321
   |||::||: : : : : : : : : : : : : : : : : :
Db 208 TYKTTGTTPBESEKTEDSRTTVASDKLLT-----KTTKNTQETISANELTQSL 254

Qy 322 KVDEEEKTEPEASEOEPAEDTDQARLSADYEKVLEPLEDOVGDFLASSEEKCAPLATE 381
   |||::||: : : : : : : : : : : : : : : : : :
Db 255 AEPTHEGGRANENTPSPAEPTENRERTAN-ENTLUSPAEPTENRRERTANENTAP---- 309

Qy 382 VFDEKMEAHGEVVAEVHVSUVTEKTEBEGGGGAEGGVWVEGTGESLPPEKLAEPPQBPVQ 441
   |||::||: : : : : : : : : : : : : : : : : :
Db 310 -FPAGPTENREMANENTTLPPAPEPTHG-----ERTANENTTPS 348

Qy 442 EAAPAELMKSRMCVSGGDHTQLTDLSPBEKTLPKHPGIGIVSEMVLMSOERI KVOGSP 501
   |||::||: : : : : : : : : : : : : : : : : :
Db 349 PAEPTEH-----GERT-----ANENTTP-----SPAEPTEHGERTPFAND- 383

Qy 502 LKKLFSSSGLKLLSGKKQKRGCGGDDERPGEVQHITHTESPSADEQKGESSASSPEPE 561
   |||::||: : : : : : : : : : : : : : : : : :
Db 384 --KTTSS-----ABSTEHERGPLANENTTPSPAEPTENRERTANENTTPSPAEPTENR 436

Qy 562 ETTCLEK---GPLEAPODGEA---EGCTTSDGKBKRG-I TPWASFKMVTPPKRVRRPS 614
   |||::||: : : : : : : : : : : : : : : : : :
Db 437 ETTANEKTLISPVEPTENRRTTANEKTTSPAEPTENGORTPANЕК-----TTSS 487

Qy 615 ESDKBEELEKVKSATLSSTDSTVSEMDQEVKTVGEE--QKPEPRKRVDTSVSWEALICV 672
   |||::||: : : : : : : : : : : : : : : : : :
Db 488 SAEPTHEGERTPLANENTTPSPAEPTENRERTANEKTTSPAEPTENGDRT----- 538

Qy 673 GSCKRKARKASSDDDEGGPPT-----LGDSHRAEASKDKKEAGTDAVPASTOEOD 723
   |||::||: : : : : : : : : : : : : : : : : :
Db 539 PLANEKTTPLSLAPTENGORTPANЕКTTSSSAEPTHEERTPLANENTTPSPAEPTENR 598

Qy 724 Q--AQGSSSPAPAGSPSEGCVSTWFSFKLVTPRKSKSLEKAEDSSVEQLSTEIEP 781
   |||::||: : : : : : : : : : : : : : : : : :
Db 599 ERTANENTTPSPAG-PPTENREMTANE--KTTLPFAEPTENR-ERTANEKTT--SSPAEP 651

Qy 782 SRBESWSYISKFFIPGRARRKADQKGEQATVEDSGPVINEDDPNVAPVPLSEYNAVE-R 840
```



```

Db 652 TEN-----GQRPFPAN--EKTSSPAEPTGEKERTPLANENTTSSPAEPTENR 697
Qy 841 EKMEAGCNTELPOLLGAVYVSELSKTLVHTVSVAVIDGTRAVTSVEBRSPMSIASYTE 900
Db 698 ERTANEKTTTPP-----AEPTEENERTANENTTP-----SPAQ 730
Qy 901 PLEHTAGAMPVEEVEVKDIIAETPVLTQTLPEGKAHDMVTSEVDFSEANTAEET 960
Db 731 PTEN--GDRPLANEKT-----TPSLAEPTENGK-----RTPFANEKTT---S 768
Qy 961 SEARTEEVTSAGAEETDMVSAVSQSLDSDPTTEEATPVQEVESGVLDFEERQTOA 1020
Db 769 SSAEPTHEAERTPLANEKT-----TSSP-----AEP-----TENRRTAN- 803
Qy 1021 ILQAVADKYKESQVPATQVORTGSKALEKV---EEVEEDSEVLASEKEKDVMPKCPV 1076
Db 804 -----EKTQFPAPTEENRE--STANEKUTPPAPTEENREWTANE--NTTSSPAEPT 852
Qy 1077 QGAGAEH-----AGSETGQATPESLEVEPVTDVHVATQVYIKLOQLMEQAVA 1127
Db 853 EHEEMTPLANEXTLSPAEPTENGERTP-----FTNEKTT 887
Qy 1128 PESSLTLDSENGSTPLADSP-----ADGTQDETTIDSODSKATAA---VRQSVTESEA 1181
Db 888 PSSAA--PTEHGERTPLANETITPPSRAPTEHGERIANEKATPSPAKTEHGETTVNED 944
Qy 1182 ATAQKEEPTLPNNVP--AOE-----EHGEERGRDYLEPTQOELTAAVPVLAKTE 1230
Db 945 TTPSAEPTENGERTPLANEKTITSPTESTEHGERTANEKTTPSAPTE----- 994
Qy 1231 VGQGEVMDLGEKYKEOEYFVHSGPNSQKADVTYDSEVGVAGCOEKESTEVQSLSL 1290
Db 995 -----GERTPSANEKTIPTSPAKTEHEMTPSA-----NEMTTSPVKP 1034
Qy 1291 EEGEMETVEKEKRTKEQVSEEGEOETAAPEH--EGTKGFVLTLDMSSRGKALGSL 1349
Db 1035 TEHGERTTLANEKITLSPBEPTEHAKTTISANEKTTPLAKTEHGERKTSNDKITSS- 1093
Qy 1350 GGSPLPDQKACIEVOYOSLDTTV---TQTAVEVEKVIETVVISSETGESPEVCAGHLL 1406
Db 1094 -AAESTHERDRATSANVITPPAPBPIKAKRTTLAHEKK--TQVTEKSTHEKTTSS-- 1147
Qy 1407 PAEKSSATGHWTLQHAEDTV-----PLGRESQAESt-----PIITPAPE 1447
Db 1148 TTEKTRTPREKPTL-YSEKTICTKGNKTPV--PEKPTENIGNTLTLETIKAVKSTENPE 1205
Qy 1448 STLHPDLOGEISASQRESEEDKPDAGDAGKSTALEKVLKAPETILEESKNKIY 1507
Db 1206 KT-----AAVTKIKPSVKYTGKSLTTTSSHLNKE--VTHQVPTGSFT 1248
Qy 1508 INVIGTAVDOFARTEAPETHAY---DSOTQVPACRLDREPNRCW 1550
Db 1249 LITSTKLSITSSEATGNESHRYLKNKDSQKGIHGMQCMENDSPFAM 1295

RESULT 9
US-09-971-536-70
; Sequence 70, Application US/09971536
; Patent No. US20020159976A1
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Havukkala, Ilka
; APPLICANT: Bloksberg, Leonard
; APPLICANT: Lubbers, Mark
; APPLICANT: Dekker, James
; APPLICANT: Christensson, Anna
; APPLICANT: Holland, Ross
; APPLICANT: O'Toole, Paul
; APPLICANT: Reid, Julian
; APPLICANT: Coolbear, Timothy
; TITLE OF INVENTION: Lactobacillus rhamnosus polynucleotides, polypeptides and Methods
; TITLE OF INVENTION: Using them

```

```

; FILE REFERENCE: 104362
; CURRENT APPLICATION NUMBER: US/09/971, 536
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: U.S. No. US20020159976A1 09/634,238
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: U.S. No. US20020159976A1 09/724,623
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: PCT/NZ01/00160
; PRIOR FILING DATE: 2001-08-08
; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 70
; LENGTH: 1879
; TYPE: PRT
; ORGANISM: Lactobacillus rhamnosus
US-09-971-536-70

Query Match          4.7%; Score 382; DB 9; Length 1879;
Best Local Similarity 19.1%; Pred. No. 1,8e-09;
Matches 324; Conservative 302; Mismatches 726; Indels 342; Gaps 70;

Qy 20 TPSELVLSGHGPAASGAGADPADADPATKLPG-KNGQLSVNGVAEQ---DVHQEE 75
Db 124 TATEKAKQKDQVAAEATKAKAIDQAQTEGCVQADAGIAIDAQHQPGLNVRREA 183
Qy 76 NQGOEEV-----VDEDVGORESEDEVREKDRVEEMAAANS-TAVEDITTKGQETSIIIE 129
Db 184 KQAIIDAEAAKVTAEIEDSTLATEKKAQKGVADDEAKAKATAID-----QAQTIIEID 237
Qy 130 Q-----IPASENVVEEV--OPAESQANDVGFKKVFK 159
Db 238 KAKDDGKAIIDAQKOGADPPTRAKQAKAIDAEAAKVAIDODPPLTAKDKTRAK-Q 295
Qy 160 FVGRFTYKKDKNEKSDPTOLLITVKKDEG--EGAEASVAGAGH---QEPSYETAVGESASK 215
Db 296 GVGBEATKAKTAIDQAKTIDGVIQAKDDGKAIIDAQHQAQGLDLTRKDSAKQAIIDAEAK 355
Qy 216 -----ESELKQSTIEKO-----EGTLKQEOSSTEIPLQABSPQAAEEBAKDE 256
Db 356 ITDAINDODTLTSTKKAQKQAVADAEAKAKVAIDQONDAI--LQAOADGKAIIDAHQ 414
Qy 257 GEEKOEKPTSPSPSPSPVNSSETTSSFKPFTGMAWGRKKTTSFKSKEDPLETAEKPK 316
Db 415 IGADLDITQKTAQAQ-----IDKEAA---KVL-----AIEDPPLTSAEK 453
Qy 317 EQAEKVDDEEKEKTEPASEEQEPAEDTDQ-ARLSAD-----YEKVELPLEDQVGDLE 368
Db 454 AQKGVADETAKATTAIDS-----ARNADELAKAQADGKAIIDAQHRLGMDLARKTKDAQ 508
Qy 369 ASSEKCAPLATEVFDEKMEAHQVVAEIVHSTVEKTEEEEO--GGGGEABEGGVVE--- 422
Db 509 AAIDAEAAKVEAIDODPPLTTSQKKAQKQTPAAEATKAKQTTIAKADADVIDIAEKAGI 568
Qy 423 -----GTGSLPPEKLAEPQEVPOEAPAEELMKSEMVCSGGDHQ---LTDLSPEE 472
Db 569 QAIIDGHQSGALLDTRKVDAAKKAIDAEBAKTNDAIIDODVLTLSAKATQKQKVTDEBAKA 628
Qy 473 KTLPKHPEGIVSEVEMLSSQERIVQSGPLKLLFPSSGLKLSGKKQKGGKGGGDEEPG 532
Db 629 KT-----AIDAKNMDVTQDAKASGIGALIDAVHQS--TLIDTRKQDAK--AIDAEAV 678
Qy 533 EY-----OHHTESPESADEQKGESSASSPEPEPTTCLKEGPLEAPODGEABEGTSSDG 587
Db 679 KVIAAIGQDVTLTQAEKTLTQQAADAVDAATQ-----AKAAIDAKNMDAVDAQKDG 729
Qy 588 EK-----KREGITPMASRKKAIVTPKRRRPSSESDKEBELEKVASA-----TLSTSTV 637
Db 730 IKAIDAQHQAQGL---ALMERKAANKL-----AETADKVAQAAGQDVTLTATQCAV 778
Qy 638 SEMODEVKTGEEQ-----KPEPRKRVDTSVSWEALICVSSKKRRARKAS 683
Db 779 QRQAITVEVTRANQALDAAGNADAVDAQKAGVKAIDQHQSGQAL-----ADRRKDAK-Q 833

```

```
Qy 684 SSDDEGGPRTLGDSH---RAEEASKKEAGTDAVPASTQORQOAGQSSSPAGSPSEG 740
Db 834 AIDAEAKETAADODATLTANAKASQOQVADEATKAKEAIDAAKQADAVDQA-----KN 889
Qy 741 EGVSTWES---FKRLVTPRKSKSLEKABDSSVEQLSTSEIPSRESWSWIKKFIPIGR 797
Db 890 DGIRAIQAHHAGQAVADRKA---AKAIDAEAKVTGNID--QDETLTATEK-----938
Qy 798 RKKRAGKQEQATVEDSGFVINEEDPNPVA--VPLSEYNVAREKWEAQG-----NTEL 851
Db 939 ----AAQKQAVATEADNAKQAIKQO--NADAVDKAKTGKGIKAIDAQHQSGQAIRQONDA 993
Qy 852 POLLAGVTVSEELSKTLVHTVSVAVID---CTRAVTSVEERSPSWISAS--VTEPLEHTAG 907
Db 994 KOAIDA--EAAKVTAIDQDDTLTAAEKKAQKQAVTDAETKAKAIDAATLVADAIDQALA 1051
Qy 908 EAMPPEVEBTEKDI--IAEBTPVLQTLPKGDHDDMTSVSDVFTSEAVTATETSEALRT 966
Db 1052 DGIKTIDAQYOTGIALDKQKAAKQTI-----DAAEAKVSEAIID--QDVTLTADQKATQK 1104
Qy 967 EEVTEASGAETTMWSAVSGLTSDPDTE-----EATPVQEVESGVLDTTEBERQTA 1020
Db 1105 QAV-----ADEATKAKAIDAQSDADAVIQATIDGIEAIDAQHQSATALD--KQKQQAQK 1157
Qy 1021 ILQAVADKVEESQVPATQTVORTGSKALEKVEVEEEDSEVL-----ASEKEKVMP 1072
Db 1158 AIDAEAKVSKAIDODVTLTATQKADQKQAVIAEADKAKKLIDAAGNADGTKQAESGK 1217
Qy 1073 KGPVQEAHGAELAQSGSETQATPESLEVPETADVHDVATQVVKLQOLMEQOAVAPESSE 1132
Db 1218 AIDAQHQSSQALADRKDAK--TAIDAEAKETAIDHDATLTA-----NEKA---SOKQ 1267
Qy 1133 TLTDSETNGSTPLADSDTADGTOQDET-----IDSQ-----DSKATAVRSQVTEE 1179
Db 1268 AVTDEATKAKKAIDAAKQADAVDQAKTDGKKAIDAQHHSGQALDDRKA---KQVIDA 1323
Qy 1180 EAA--TAQKEPSTLPNNVPAQEEHGPGRDVLPTQOELTAAAVPVLAKTEVQGEDEV 1237
Db 1324 EAAKVTAIDQDNTLTAKAQAQKQGVATEAD-----KAKQAIDAAGDA 1367
Qy 1238 DMLDGEKVEQEVFVHSGPNKQAAADVTYDSEVMGVAGQEKESTEVQSLSLEEGEMET 1297
Db 1368 DAVDQAKTAGIAI-----DAQHAGKTIUSRHDDAKQAIDSEAAKV-----IKAIDODP 1417
Qy 1298 DVEKEKETPEQVSEGEQETAAPHEG-----TYG-----KPVLTLDMPSS 1340
Db 1418 TLTAQKEAQKQAVATEADKAKKAIDAGDADAVDQAKTAGIKAIIDBQHKSGQIVDARKE 1477
Qy 1341 ERGKALSGSGSPSLPDQKAGCIEVQVSLDVTVTQTAAEAVKVIETVVISGETGESPEC 1400
Db 1478 DAKKAIDAEAGKVTDAIDHDATLTAAQKEAQKQAVADEADKAKKAIDAAGNADAIDQAKS 1537
Qy 1401 VGAHLLPAEKSSATGGHWTLOHAEDTVPLGPESQAESPIIIVTAPESTLHPDLOGELISA 1460
Db 1538 AGIKAIIDBQHKSGQ-----SIDTRKDDAKKAIDGVEAKITDAIDHDPTLT-----DAEKAT 1588
Qy 1461 SQRESSEEDKPDAGPDAGDKESTAIEKVLKAEPEILELESKNK---IVLNVITQAVDQ 1517
Db 1589 QKQAVIAEADKAKKAIDAAAG--DADAVDQAKGAKIKAIIDQHQKSGQALAIRKDAKKAIDE 1647
Qy 1518 PARTETAPETHAYD 1531
Db 1648 ----EAAKVSEAIID 1657
```

RESULT 10

```
US-09-815-242-5816
; Sequence 5816, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
```

```
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
```

```
; FILE REFERENCE: ELITRA.011A
```

```
; CURRENT APPLICATION NUMBER: US/09/815,242
```

```
; CURRENT FILING DATE: 2001-03-21
```

```
; PRIOR APPLICATION NUMBER: 60/191,078
```

```
; PRIOR FILING DATE: 2000-03-21
```

```
; PRIOR APPLICATION NUMBER: 60/206,848
```

```
; PRIOR FILING DATE: 2000-05-23
```

```
; PRIOR APPLICATION NUMBER: 60/207,727
```

```
; PRIOR FILING DATE: 2000-05-26
```

```
; PRIOR APPLICATION NUMBER: 60/242,578
```

```
; PRIOR FILING DATE: 2000-10-23
```

```
; PRIOR APPLICATION NUMBER: 60/253,625
```

```
; PRIOR FILING DATE: 2000-11-27
```

```
; PRIOR APPLICATION NUMBER: 60/257,931
```

```
; PRIOR FILING DATE: 2000-12-22
```

```
; PRIOR APPLICATION NUMBER: 60/269,308
```

```
; PRIOR FILING DATE: 2001-02-16
```

```
; NUMBER OF SEQ ID NOS: 14110
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 5816
```

```
; LENGTH: 2478
```

```
; TYPE: PRT
```

```
; ORGANISM: Staphylococcus aureus
```

```
US-09-815-242-5816
```

```
Query Match 4.7%; Score 377; DB 10; Length 2478;
```

```
Best Local Similarity 18.3%; Pred. No. 4.1e-09;
```

```
Matches 305; Conservative 274; Mismatches 625; Indels 462; Gaps 68;
```

```
Qy 47 PATKLPKQKQSLSVNGVABQGDVHVEENQEQEE--EVVDDEVQGESESDVREKD----101
```

```
Db 681 PTVKPQAKQDIIQAVTTRKQKQ---IKKSNASLQDEKDVANDKIGKTIETRAIKDAIDAAAT 736
```

```
Qy 102 RVEEMANSTAVEDITKDGQEQETSEIIQIOPASENNVEEMVQPAESQANDVGFKKPKFV 161
```

```
Db 737 NAQVEAIKTKAINDINTTTPATTAK-----AAALEEFDEVVQAQIDQA-----779
```

```
Qy 162 GFKFTVKKDKNEKSDTVQLLTVRKDEG--EGAEASVAGDGHQEPSVETAVGESASSELK 220
```

```
Db 780 --PLNPPTTNEEVAEAIERINAAKVSQKAIETTTAQD-----LERVKNEEISKIENIT 832
```

```
Qy 221 QSTE-----KQEGTLKQESST-----EIPLOAESDQAAEBEAKDEG-----257
```

```
Db 833 DSITQTKMDAYNEVKQAATARKAQNAVTVSNATNEEV---AEAD-AAVDAAQKQGLHDIQV 888
```

```
Qy 258 EEQKEKEPTKSPSPSPVNSSETTSPFKKFTTHGAWGRKKTSPFKSKEDDLETAERKE 317
```

```
Db 889 KSKQEVADTKS--KVLDKINAIQT-----QAKVPAADTEVENAYNTRK 930
```

```
Qy 318 QE-----AEKVDBEEKEKTPASBEOQPAEDTDOARLSADYKVELPLEDQVQGLEASSE 372
```

```
Db 931 QETQNSNASTTEEKQAAYTDLTKKQEARINLDAANTNSD-----970
```

```
Qy 373 EKCAPLATEVDEKMEAHQEVVAEVHVSVTEKTEEOGGGGEAGGVVVGSTGSLPPEK 432
```

```
Db 971 -----VTTAKDNSIAAINQVQAATTKKSDAKAEIAQKA---SERKTAIEAMNDSTTEEQ 1021
```

```
Qy 433 LAEPQEVPOEAEP-----BELMKSRMCVSGGDHTQLTD-----LSPKETLPKHEGI 482
```

```
Db 1022 QAAKDKVDQAVVTANADIDNAAANNDV-----DNAKTTNEATTAATPDANVPAKQAI 1076
```

```
Qy 483 VSEVMLSQOERIKVQSGPLKFLSSGLKSLGKKGKGGGGDEEPGEYQHITESP 542
```

```
Db 1077 ADKVQ-----AQE-----TADGNNGSTTEEKAAKQOVQTEK--1109
```

```
Qy 543 ESADEQKGESSASSPPEPETTTCLEKPLEAPQDGEAEGTTSDEKREKREKGPWASFKK 602
```

```

Db 1110 TTADAIDAHTNAEVAKAKAAIAK--IEAIQPAT-----TTKD-----NAKE 1151
QY 603 MWTPEKVRPESDKEEELKVKSAITLSIDSTVSEMODEVKTVGEOKPEEPRKRVDT 662
Db 1152 AIAATKAMERKATIAIQTQGITAEETIAANAVADVNATQANSIIEAANSQNDVQAKTGEN 1211
QY 663 SVSWEALICVSSSKKRAKASS-----DDEGPRLLGDSHRAEASADKEKGTD-- 713
Db 1212 SI--DQVPTYNKAKATANEITAILNNKLOEIQATP-----DATDEEQADAEANTENG 1264
QY 714 ----AVPAST--OEODQAGSSPEPAGSPESGEVSTWESFKRLVTPR--KKSXKLE- 764
Db 1265 KANQHISATTNAGVDEKAKANA-----EAIINAVTKVYKKAQAKBEI 1307
QY 765 EKADDSVEQLSTELEPESRESWSIKKIPGRKKRADGKQEQATVEDSGFVEINEDDP 824
Db 1308 DQLOATQTNVINNDQNAATTEKEAII-----QQLATA-----VTDAKN 1345
QY 825 NVPAVPLSENNAVEREKMEAGNTELPQLGAVYVSELSKTLVHTYSVAVIDQTRAVT 884
Db 1346 NITAA--TDDNGVDQADAGKNSIQSTQPATVAKSNKANDVDQAVTQONQAIIDTTGAT 1402
QY 885 SVEERSPEMISASTEPLEHTAGEAMPVEEYTEKDIIAETPTVLTQTLPEKCD-AHDDM 943
Db 1403 TEE-----KNAKDLVLKAKKAYODILMAQT--TNDVTQIKQDAVADI 1444
QY 944 VTSEVDPTSEAVTATESEALRTEEVEASGAETTDVMSAV-----SOLTDSPTTEEA 998
Db 1445 QGITADTTIKVAVKDELATKANEGKALLAQPADATTEKEQONQVDQQLTGONINENA 1504
QY 999 TPVQEVESGVLDTSEETQTAI--LQAVADKVESQVPATQVQRTQSGKALEVEEVE 1056
Db 1505 QSIDDVNTA-----KDMAIQAIIDIQASTD-VKTNARAEELTEQO-----NKITEIL 1550
QY 1057 EDSEVLASEKEKDVMPKGPVQAGAEHLAQSEGTQATPESLEVEPVATVDVHATCOVI 1116
Db 1551 NNNETTNEEKGNDI--GPRVRAVEEG-----NNINATTTGIDVTATADTAVQ 1596
QY 1117 KLQOLMEQAV-APESSETLTSETNGSTPLADSDTADGTQOD-----ETIDSQDSKATPA 1170
Db 1597 KVQQLHANPVKKPKAGKELDQAAADKKTQI--EQTPNKSQOEINDAKQEVDELINQAKTN 1654
QY 1171 VRQSCVTEEPATQAKPEPSTLNNVPAOEHGEHPGRDVLPTQOELTAAVPLVLAKE 1230
Db 1655 VDQSS-TNEVYDVNAVKKEKAKI-NAVKTFSEYKCD-----ALAKIE 1693
QY 1231 VGOEVEVWMLGKEV-----KEBOEVFVHSG----- 1256
Db 1694 DAYNAKVADNSNASTSEIAEAKQIAELKQIADQNVNQAITSKODILEVQIHNLDNIN 1753
QY 1257 ----PNSOKAA-----DVTYDSEVMGVAQOEKESTEVQSL-- 1288
Db 1754 DYTTPGKKEGATIDLVAVADQKKNISADTNATQDEKQQAIIKQVDQNVQTALESINNGV 1813
QY 1289 -----SLEGEKETTVEKEKRETKPE-----QVSEGEQETAAPEHGTGYKQVLLD 1336
Db 1814 DNGVDVADLTQGAIDAIDQVATVPRANQAIEVKAEIDTESIOSQD----- 1862
QY 1337 MPSSRGAALSLGSGSPPLPDQDKAGCIEVQVQSJDTVTQTAFAVEK-----VIETV 1389
Db 1863 LTAERKTALMTI--KOITDQAKQI-----TDATTTAB-VERAKAQGLEAFDIT 1909
QY 1390 VISEGESPECVGAHLIPAEKSSATGCHWTLQHAEDTVPLG-----PESQABSI 1438
Db 1910 QIDST-----EKOKALE--ELEFALDQIEAGVNNVADATTEKEKFTNAL 1952
QY 1439 PIIVTAPBESTLHPDLOEISASQRESEEDKDPAGPDADKESTALEKVKAEPEIL- 1497
Db 1953 EDLSKATEDISDQTTNAEIAITVKNSALEQLKQRIINPEVKNALEAIREVVNKQIEIITK 2012
QY 1498 --ELESKNKIVLVNIQTAVNDQPARTETAPETHAIDSQO--VA 1538

```

```

Db 2013 NADADASAKETIARTDLGRYFDRFADKLDKTQTNVAEVALQNVTI PA 2058

RESULT 11
US-09-815-242-12967
Sequence 12967, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 1410
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12967
LENGTH: 2478
TYPE: PRF
ORGANISM: Staphylococcus aureus
US-09-815-242-12967

Query Match 4.7%; Score 377; DB 10; Length 2478;
Beat Local Similarity 18.3%; Pred. No. 4.1e-09;
Matches 305; Conservative 274; Mismatches 625; Indels 462; Gaps 68;

QY 47 PATKLPRKNGQLSSVNGVAEGDVHVOENQGOE-EVVDVEDVGOSESEPREKCD----- 101
Db 681 PTVPQAKQDIIQAVTTKQD---IKSNASLDQDEKDVANDKIGKIEITKAIKIDATTT 736
QY 102 RVEEMANSTAVEDITKQGOETSEIIEQIPASENNVEENVQPAESQANDVGFKVFKFV 161
Db 737 NAQVEAIIKTAINDINQTPATTAK-----AAALEPFDEVVQAQIDQ----- 779
QY 162 GFKPTVKKDKRKESDVTQQLTVKKDEG-EGAEASVAGDHPBEVETRAVGSASKESEFLK 220
Db 780 --PLNPDTTNEEVAEAIERIRIAAVSVGKALEATTATQD-----LERKQNEISIKENIT 832
QY 221 QSTE-----KOEGLTKOESQST-----ELIPQAESDQAAEEBAKXG----- 257
Db 833 DSTQTKDAVNAVEVQKATARKAQNAATVSNATNEEV--AEAD-AAVDAAQOGIHLDIQV 888
QY 258 EEKQKEPTKSPESPSPVNSSETTSFKKFTTHGAWGRKKTSPKSKEDDLETAERKKE 317
Db 889 KSKQEVADTKS--KVLQKINAIQT-----QAKYKPADVEVENAYNTRK 930
QY 318 QE-----AEKVDEEKEKTEPASEQEPADPTDQARLSADYEKVELPLEDQVGLBASSE 372
Db 931 QEINSNASTTEEKQAVTEIDTKQKEARTNLDANTNSD----- 970
QY 373 EKCAPLAEVVEDEKMEHQAQEVVAEVHNVSTVEKTEEEOGGGGEAGGVVVEGTGSELPPEK 432

```

Db 971 -----VTTAKDNSIAAINOVQAATTKSDAKAEIAQKA-----SERKTAIEAMNDSTTEEQ 1021
Qy 433 LAEPQEVPOEABPA-----EELMKSRMCVSGDHTQLTD-----LSPKEKTLPKHPGEG 482
Db 1022 QAAKDKVDQAVVTANADIDNAANDV-----DNAKTTNEATIAAITPDANVVKPAQAQAI 1076
Qy 483 VSEVEMLSOBERIKVQGSPLKLFSSSLGKLKSGKKQKGGGGGDEPGEYQIHHTESP 542
Db 1077 ADKVQ-----AQE-----TAIDGNGSTTEEKAAAKQOVQTEK- 1109
Qy 543 ESADQEGESSASSPEPEETTCLEKGPLEAPQDGEABEGTTSDEGKREGITPWASFKK 602
Db 1110 TTADAAIDAATHNAEVAEAKKAATAK--IEAIQPAT-----TTKD-----NAKE 1151
Qy 603 MVTPEKVRPPSSDKKEEELKVKASATLSSTDSTVSEMQDEVKTVGEEQKPEEPKRRVDT 662
Db 1152 AIATKANERKTAIAQTODITAEIAAANADVNAVTAQANSIEAANSONDVDAQTGTEN 1211
Qy 663 SVSWEALICVSSKKRARKASS-----DDEGPRTLGDSHRAEASKDKEAGTD-- 713
Db 1212 SI---DQVPTVVKKATARNETAILNKKLOEIQATP-----DATDEKQAAADAENTENG 1264
Qy 714 ----AVPAST--QPDQAQSSSPPEAGSPGEGVSTWESFKRLVTPR--KKSXSKLE- 764
Db 1265 KANQAISSAATNAQVDEAKANA-----EAINAVTPKVVKQAQKDEI 1307
Qy 765 EKABDSSVEQLSTIEIPRSBESWISIKFIPGRKKRADGQEOATVEDSGVINEBDDP 824
Db 1308 DQLOATQTNVINNDQATTEKEAAI-----QLATA-----VTDKKN 1345
Qy 825 NPAVAVPLSENAVEREKMEAGQNTPELLGAVVYVSEELSKTLVHTVSVAVIDGTRAVT 884
Db 1346 NITAA-----TDDNGVDQAKDAGKNSIQSTQATPAVKSNAKNDVDOAVTTQNOAIDNTTGT 1402
Qy 885 SVEERSPSWISASVTEPLEHTAGAMPVEEVTEBKDIIAETPTVLTQTLPEGKD-AHDDM 943
Db 1403 TEE-----KNAKDLVLKAKKAYQDILNAGT---TNDVTQIKDOQAVADI 1444
Qy 944 VTSEVDTSEAVTATSEALRTEVEVTEASGAETTMVSAV-----SOLTDSPTTTEEA 998
Db 1445 QGITADTTIKDVADELATKANEOKALIAQTADATTEKEQANQOVDAQLTQGNQNIENA 1504
Qy 999 TPQVEVSGVLDTTEERQTOAI--LQAVADKVKKEEQVPATQTVQRTGSKALEKVEVE 1056
Db 1505 QSIDDVNNTA-----KONAIQAIQDIOASTD-VKTNARAELLTMOQ-----NKITEIL 1550
Qy 1057 EDSVLASEKEKVMKPGPVQEGAGAEHLAQSGSETQATPESLEVPETADVHDVATQCVI 1116
Db 1551 NNNETTNEEKNDI--GPVRAAYEEGL-----NNINAATTGDTVTTAKDTAVQ 1596
Qy 1117 KLOQLMEQAV-APESSETLTDSETNGSTPLADSDTADGTQDQ-----ETIDSODSKATAA 1170
Db 1597 KVOQLHANPVKKPAGKXELDQAAADKKTQI--EQTPNASQOEINDAKOEVDVTELNOAKTN 1654
Qy 1171 VRQSVTTEERAAQKEPSTLPNNVPAQEBHGEPPGRDVLPTQOELTAAAVPVLAKTE 1230
Db 1655 VDQSS--TNEYVDNAVKEGKAKI-NAVTFSEYKKD-----ALAKIE 1693
Qy 1231 VQGEVGDWLDGKGV-----KEQEVFVHSG----- 1256
Db 1694 DAYNAKVNEDNSNASTSSEIAEAKQKLAELKQTDQNVNQATSKDDIEVOIHNLDLONIN 1753
Qy 1257 ----PNSQKAA-----DVTYDSEVMGVAGQCKEKESTEVSQSL----- 1288
Db 1754 DYTPTCKESATDLVAYADQKNNISADTNATQDEKQQAQKQVDQNVQTALESINNGV 1813
Qy 1289 -----SLREGENETDVEKEKRETKPE-----QVSREGEQETAAPHEGTYGKPVLTLD 1336
Db 1814 DNGVDVDDALTOGKAADIAIQVDAVTKPKANOAIKVAEDTKESIDQSDQ----- 1862
Qy 1337 MPSSERKALGSLGGSFSLPDQDAGCIEVQVQSLDVTVTOTAEAVEK-----VLETV 1389
Db 1863 LTAEKTEALAMI---KQITDOAKQGI-----TDATTAB-VEKAKAQGLEAFDNI 1909

Qy 1390 VISGTGSPSCVGAHLLPAKSSSATGHHWTLOHABDTPVLG-----PESQAESI 1438
Db 1910 QIDST-----BKQKAIE---ELETALDQIEAGVNVNADATTEEKEAFTNAL 1952
Qy 1439 PIIVTPAPESTLHPDLOGEISASQRESSEEDKPDGPDADGKRESTAIEKVLKAEPEIL- 1497
Db 1953 EDILSKATEDISQTTNAETATVKNSALEQLKQAKRINPEVKKNALAIREVNVKQIELIK 2012
Qy 1498 --ELESKNKIVLNVIOQTAVDQFARTETAPETHAYDSQTQ---VPA 1538
Db 2013 NADADASAKAIEARTDLGRYDFRADKLDKTKTNAEVAELQNVITPA 2058
RESULT 12
US-09-919-172-98
; Sequence 98, Application US/09919172
; Patent No. US20020119463A1
; GENERAL INFORMATION:
; APPLICANT: Faris, Mary
; APPLICANT: Turner, Christopher M.
; TITLE OF INVENTION: PROSTATE CANCER MARKERS
; FILE REFERENCE: PA-0036 US
; CURRENT APPLICATION NUMBER: US/09/919,172
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/222,469
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PERL Program
; SEQ ID NO 98
; LENGTH: 3256
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020119463A1 2700132CD1
US-09-919-172-98

Query Match 4.7%; Score 377; DB 10; Length 3256;
Best Local Similarity 20.2%; Pred. No. 5.6e-09;
Matches 387; Conservative 241; Mismatches 762; Indels 524; Gaps 83;

Qy 14 QPAGSDTPSELVLSGH-----GPAEASGAGDPADADPATKLPQKNGQL 58
Db 1412 QTGETTHTDKVPGEKDSINAFRETAQKLDPAASVTSGRHP-----KTKKAQPL 1464
Qy 59 SSVNGVAEQGDVHVQENQEQEE-----EVVDEDVQGR--ESEDVREKDRVEEM 106
Db 1465 EDLAGWKELFQTPVCTDKPTTHEKTKIACRSQDPDPVPTTSSKPKRSURKRVDBEEF 1524
Qy 107 AANSTAVEDITK-----DQOE-----ETSEIIEQIPASENNVEENVQPAES 147
Db 1525 FALKRTPSAGKAMHTPKPAVSGEKNIYAFMGTPVQKLDLTENLTGSKRRLQTPKEKAQA 1584
Qy 148 QANDVGKVKFVGFKFTVKDKNEKSDTVQLLTVKKDEGEAEAS-----VGA 197
Db 1585 LEDLAGFKELFQTRG--HTESMTNDKTAKVACKSQDLDKNPASSKRLKTSILGKVG 1642
Qy 198 GDHQEPSVETAVGESASKESELKQSTEQKGTLEKQESSTEIPLQAESDQAAEEAKDEG 257
Db 1643 KEEL-----LAVGKLQTSGETTHTEPTGDKGSKMAFMESP-----KQILDSAAALTG 1692
Qy 258 EEKQKEPTKSPSPSPVNV---SETTSSFKKPFTHGWAGWRKKTSPKSKESDDELTAE 313
Db 1693 SKQLRTPKSKSEVPEDLAGFIELFQTPSHTKESMTNEKT---TKVSYRASQPLVDTP 1749
Qy 314 KRKEQ--EAEKVDEEEK---EKTEPAS-----EBOE-----PAEDTDQ- 346
Db 1750 SSKPQPKRSURKADTEEEFLAFKQTPSAGKAMHTPKPAVSGEEKDINTFLGTVPQKLDQ 1809
Qy 347 -----ARLSADYEKVELPLEDOVGDL-----ASSEK-----CAPLATBEVD 384
Db 1810 GNLFSGNRRLOTRKEKAQ-ALEELTGREFLQTPCTDNPTTDEKTKKILCKSPQSDPAD 1868

```

QY 385 -----EKMENQEVVA--EVHVSVEKTEEEQGGGAEAGVVEGTGSLP 429
D 1869 TPINTKORPKRLKKAADVEEFLEFKLTPSAGKAMHTPKAAGEKQINTFVGI-----P 1924
QY 430 PEKX-----AEPOVEQAEPAELMKSRMCVSGDHQO-----LTFDS--- 469
D 1925 VEKLDLGNLPGSKRRPOTPEKAKALEDLAGFKLFQTPG--HTBESMTDDKITEVSKS 1983
QY 470 --PEEKTLPKHREGIVSEVEMLSQERIKVOGSP--KLFSSSGKLKSGKQKGRG 524
D 1984 PQDPVPKTP-----TSSKQRLKISLQKVGKVEEVLVPGKLTQTSKKTQTHRB 2031
QY 525 GGGDEPEEYOHITESPESA--DEQGESSASSPEEP--EETTCLE-----KGPLEAOD 576
D 2032 TAGD-----GKSIKAFKESAKQMLDPANVGTGMEWRPTPEKQSLDLAGELEFOTPD- 2087
QY 577 GAEBSGTTSDGKKREGITPMASFQKMTVPKKRVRRP-----SESDKEELEKVKSATLS 632
D 2088 -HTBESITDDTKTKIACKSP--PPESMDTPSTRRRPKTPLGKRIIVELSMLKQLTQTT 2144
QY 633 -TD-----STVSEMODEVKTV--GEQKPEPKRVDTSVSWEAL----- 669
D 2145 HTDKVPGDEKGINVFRFPAKQKLPAAVSATGSKRQPRTPKGAQPLEDLAKELFQTP 2204
QY 670 ICVG-----SSKKARAKASSDDGCGPTLAGDSHRA-----EASQKDKAGTDAY-- 715
D 2205 ICTDKPTTHEKTKIACKSPQDPDPVGTPTIFPKQSKRSRKADVEEESLALRKRTPSVGK 2264
QY 716 -----PASTOQODQAQSSSP--EPAGSPSEEGVSTWESFKRLVTPRKSKSKLEEKAE 768
D 2265 AMDTPKPPGGBDKMAKMGTPVQKLDLPGNLPGSKRMPQ-----TPKEXQA--LEDLAG 2318
QY 769 DSSVEQLSTELPSREESVSI-----KKFIPGRKKRQADGQEQATVED 813
D 2319 FKELFQGTGKTPTDEKTKIACKSPQDPDVDTASTKQPKRMRLRADVVEEFALRK 2378
QY 814 SGPEINDDPNVAVVPLSEYNA--VER--EKMENQGN-----TELPGLLAGV 859
D 2379 RTPSAGKAMDTPKPAVSDEKQINTFVEETPVQKLDLGNLPGSKROPQTPPEKAEALIEDLV 2438
QY 860 VSEELSKTLVHTVSAVVIDGTRAVTSVEERSPSWISASVTEPLEHTAGAMPVEEVTEK 919
D 2439 GFKELPOTPGHTEESMTD--KITEVGSCKSP-----QESPKTSASSQRLKIPLV 2487
QY 920 DIIAETTV-----LTOTLPEGDADHD-----MWTSEVDTSEAVTATETSEALR 965
D 2488 KYDMKEEFLAVSKLRTSGEITQHTHEPTGDSKSIKAFKESPKQILDPAASVATGSRQRLR 2547
QY 966 T-----EYVTEASGAEF--TTDMVASVOLTUSPTTEATVQEVESGV 1008
D 2548 TRKEKARALEDLVDPEKELFSAPAGHTEESMTIDKNTKIPCKSPPELITDPAATSKCPKTR 2607
QY 1009 LQTEEEBQOTQAILQAVADKVEESQVPAQTQVQRTGSKALKVEEVEDSVLASEKEX 1068
D 2608 LRKEKKEE--LSAVERLTQTSOSTHKEKPPASGDEGIVKQRAKKKNNPVEEBS 2662
QY 1069 DVMKGPVQEA--GAEHLA--QGESGQATPES-----LEVEVTADVDA 1109
D 2663 RRRPAPRPEKAQPLEDLAGFTLSETSHTGSLAGKATKIPCSBPLEVVDVTASTRK 2722
QY 1110 VATCOVILKQOMEQAVAPESSETLTDSETNGSTPLADSDTA----- 1151
D 2723 HURTVQKVO-----VKEEPSAVKFTQTSGETTDADKEPAGEDKGIKALKESAKQTPA 2775
QY 1152 -----DGTQOQDETI-----DSQDSKATTAARQSQVTEEBAA 1182
D 2776 PAASVTGRRRRAPRESAQALIEDLAGFKDPAAGHTEESMTDKTKIPCKSPPELIEDTA 2835
QY 1183 TAOKEEPTLPPNNVPAQEEHGEEPGRDVLPTQOELTAAAVPVLAKTEVGOGEVDMLDG 1242
D 2836 TSSKRRPTRAKQVKEE-----LLAVKLTQOT-----SGETTHMDX 2873

```

```

QY 1243 EKVKEOEVFVHSGPNSOKADAVTYDSE--VMG-----VAGQEKESTE 1284
D 2874 EPVEGKGTAKFKQPARK-----LDABDVIGSRQRRAPREKQAPLEDLASFELSQTP 2928
QY 1285 VQSLSEEGEME--TDVEKEKRET--KPEOVSEEGEOETAAPENHGTGKPYLTJLDMESSE 1341
D 2929 GHTBELANGADSDPTSAFKQTPDGSGLPKIS---RVLRAPKVE--PVGDVSTRD--PVKS 2983
QY 1342 RGRALGSLGSPSLPDDQKACIEVQVQSLDTTYQTQAE-----AVEKVIETVVISERGE 1396
D 2984 QSKNTSL---PPLPFRGGG-----KQSVGTGKRLCWPAPBEIYEE----- 3024
QY 1397 SPECGAHLIPAEKSSATGGMHTLOHAEDEVPLQAPESQAESIPITVPAPSTLAPDLOG 1456
D 3025 -----LPASIK-----QRAVAPRAGKSSSEVVIMKSLRISAKRIEPAE 3063
QY 1457 EISASORESEEDK--PDAGDADG--KESTALIEKYLAEPEILELESKNKIVANVIOQA 1514
D 3064 ELNSNDMKTNEEHKLODSVENKGISILRSRQKTEAEQOITEVFVLAERIEIN----- 3118
QY 1515 VDQFARTE-----TAPETHAV--DSQTVPACRLDSREPNRCWTRMDAKKHP 1561
D 3119 -----RNEKKEPKTSPEMDIQNPDDGARKPIPRDKVTENKRCILRSARQNESSOP 3167

```

```

RESULT 13
US-09-764-176-7
; Sequence 7, Application US/09764176
; Patent No. US20020127553A1
; GENERAL INFORMATION:
; APPLICANT: NOTEBORN, Mathieu Hubertus Maria
; APPLICANT: DANEN-VAN OORSCHOT, Astrid Adriana Anna Maria
; APPLICANT: ROHN, Jennifer Leigh
; APPLICANT: WEISS, Bertram
; TITLE OF INVENTION: APOPTIN-ASSOCIATING PROTEIN
; FILE REFERENCE: 4725US
; CURRENT APPLICATION NUMBER: US/09/764,176
; CURRENT FILING DATE: 2001-01-17
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 1400
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Description of Sequence: Amino acid sequence deduced from the nuc
US-09-764-176-7

```

```

Query Match 4.6%; Score 368; DB 10; Length 1400;
Best Local Similarity 19.6%; Pred. No. 5.2e-09;
Matches 283; Conservative 226; Mismatches 495; Indels 442; Gaps 65;

QY 95 BDVREKDVEMANASTAVEDI--TKQGEETSEIIEQIPASENNVEWQAPESQANDV 152
D 92 DNLKFNKINLEDADHTMLQPIGRDKGLMTWYQLDQ-----DHNVNRYIEQDDQDS- 145
QY 153 GFKKVFEVGRFTVKDKDNKSDPTVOLITVK-----KDEGEAGASVAGADHOEPSVE 206
D 146 -----SWKCIY--RNRNELAETLALLKAQIDPVLLKNSQQONS-----RSPLE 190
QY 207 TAVGESAKSEBELKOSTEKQGTCLKQEOSSSTEIPLQAESDQAAE--EAKQDEGEKQKE 264
D 191 ---DEETKKEETPKQEOQKE---SEKKKSEEOPMWLENRSTANVLEETTVKKEKEDKE 244
QY 265 PTK-----SPESSSPVNSSETTSSFKKFFTHGAGARKKTSFKSKSEDDLET--- 311
D 245 LVKLPLVILKELPLENEKKTIKEESDSFKENYK-----PIKVEVKECRADPKOTSS 298
QY 312 AEKKEQEAQKVD-----EEKEKTEPASEQEPALDTQD-----RLSADYE--- 354
D 299 MEKPAQEPERIEFGGNKSKSHETKSTETETKLNQDQAKIPLKKEKILSDDFDPSV 358

```

QY 355 -----KVELPLEDOVGLERASSBEKCAPLAT-----EVPDEKMEAH----- 390
Db 359 KGPLCKSVTPTKFELKDEIKOEETCKRIISTITAGHEGKQLVNGEVSDEKVAENFKTEP 418
QY 391 -----QVWAEVHVSTVEKTEEOGGGGEABGGVVVE-GTG-----ESLPPEK----- 432
Db 419 IETKFYETKEESYPSKORNIIITNGTGESLNSVTSMKGTGELKETAPURKODADSSISV 478
QY 433 -----LAEOVEPQE-----AEPAEEL-----MKSREMCVSGGDH----- 462
Db 479 LEIHSOKAQIEEPPPEMETSILDSSEMAKOLSSKTALSSCTMKGEKSPKTKDKRXP 538
QY 463 -----TQLTDLSPBEKTLPK-----HPE 480
Db 539 PILECLEKLEKSKTFLDKAQLSPISPEEVPKSTLSEKPGSPAAETSPPSNIIDHCE 598
QY 481 GIYSEVEMLSQERIKVQGSPLKLFSSGLKLGKQKQKGGGGDEBPGEVQHIIHE 540
Db 599 KLASEKEVVCQSTSTVGGQSVKKV-----DUETL-----KEDSEFTKVEMD 640
QY 541 SPESADKQGEASSASSPEEPTTCLBKPLE-----APQDGEABEGTTSDEKKR 591
Db 641 NLDNA-----QTSGEIEPSET-----KGSQKSKFYKLVPEEETTASENTEITSERQK 689
QY 592 EGI-----TPWASFVKMWTPKRVRRP-----SESDKEELEKV-----KSATLSSTDSTVSE 639
Db 690 EGILATIRISRRKKPDPSPKVPLEPENKQBEKTEKEETNVGRTLRRSPRISPTAKVAE 749
QY 640 MODE-----VKTVGEEOKPEE-----PKRRVDTS-----VSWEALIC 671
Db 750 IRDQADKKGEGEBEVEESTALQTKDKKILKSEKDTNSKVSKPKGKVRWT----- 805
QY 672 VGSKKRKAASSDDGGGPRTLGGDSHRAEASKEAGTDAVPASTOBODAOAGSSSP 731
Db 806 --GSRTEGRWKYSSNDE-----SESGSEKSSAASEEE-----EKSEEAAILADDD 850
QY 732 EP---AGSPGEGGVSWESFKR-----LVTP-----RKSKSKLEKA 767
Db 851 EPCKKCLPNHPILLICDSGTYHTACLRLPPLMIIPDGEMFCPPCQHKLLCEKLEEQ 910
QY 768 EDSSVEQLSTELPSPRESVW---SIKKETP-----GRKKRADGQOQATVEDSG 815
Db 911 QDLVDALKKEBARERKRLVYVIGISLENIPQEPDFSEDBEKKDKSKKANLERR 970
QY 816 PV-----EINEDDPNVAVPLSEYNAREKMEAOQNTBLPQLLGAVVYVEELSKT 867
Db 971 STRTRKICISYRFDFAID-----EAIEDDIKEADG-----GGVGRGKOISTI 1014
QY 868 LVH---TVSVAVIDGTRAVTSVEERSPSWISAVTEP-----LEHTAGRAMPP 912
Db 1015 TGHKGDIS-TILDEERK-----ENKRPQAAAAARRKKRRRLNDLSDSNLDEESEDPEK 1069
QY 913 VEEVTEKO-LIAEETPVLTQLEPGKADHDMVTSEVDFTSEAVTATETSEALRTBEVTE 971
Db 1070 ISGSDQEFVVDENPDESSEDPPSND-----DSDTDFCSRRLRRHPSRPMQSRRLR 1123
QY 972 ASGAETTDMSAVSOLTDSPDTTEATPVQEVESGVLDT-----EEEBRQTQAILQAVA 1026
Db 1124 KTPKKYSD-----DDEESEEENS---RDSSEPSDDFSDDFVETRRRRRRNRKQK 1173
QY 1027 DKVKESQVPATQTVQRTGSKALEKVEEVEDSEVLASEKEDVMPKGPV-QBAGAEHLA 1085
Db 1174 INYKEDSESDGSKSLRRG-KEIRRVHK-----RRLSSSESESYLSKNSDEDELAKESKR 1228
QY 1086 QGSETGQATPESLEVPVADVDHATCQVILQQLMEQAVAPESSETLTDSTNGSTPL 1145
Db 1229 SVKRGKSTDEYSEADDEEEEBEG-----KPSRKLHRIETDEEBCDNAHGDAQPA 1281
QY 1146 ADSDTADGTQOQDITSDSKATAAVRQSQVTEEEAATAQK-----EFPSTLPNN 1195
Db 1282 RDS-----QPRVLFSEQUESTKKPYRIESDEBEDFENVKGSPLDYSLVDLPST---- 1330

QY 1196 VPAQEBHGEPEGRDVL-----EPTQQ-----ELTAAAVPVLAKEVQGEVDMLDGEKV 1245
Db 1331 -----NGSPGKRAIENLIGKPTKSTPKDNSTAGSLASNGTSGQEA-----GAPE 1378
QY 1246 KERQEV 1251
Db 1379 EEEDEL 1384
RESULT 14
US-09-963-875-1
; Sequence 1, Application US/09963875
; Patent No. US20020164307A1
; GENERAL INFORMATION:
; APPLICANT: Massachusetts General Hospital
; TITLE OF INVENTION: Stem Cells of the Islets of Langerhans and Their Use in Treating
; TITLE OF INVENTION: Mellitus
; FILE REFERENCE: 17633/1235
; CURRENT APPLICATION NUMBER: US/09/963,875
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US60/169082
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: US 60/215109
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: US 60/238880
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 09/731261
; PRIOR FILING DATE: 2000-12-06
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1
; LENGTH: 1618
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-963-875-1
Query Match 4.4%; Score 354; DB 9; Length 1618;
Best Local Similarity 20.1%; Pred. No. 2.5e-08;
Matches 342; Conservative 224; Mismatches 623; Indels 510; Gaps 75;
QY 183 VKKDEGEAEASVAGDHPQPSVETAVGESASKE-----SELKQS--T 223
Db 26 VKALEEQNBLLSAGLGLRRQSADTWSRAHADDELAALRALVDQRWREKHAAEVARDNLA 85
QY 224 EKOEGT-----LKQEOSSTEI-----PQAE-----SDQAAEEBAKDS----- 256
Db 86 EELEGVAGCEQLRLARETTTEVARNRRAVEAKCARAWLSQGAELERELEALRAVHE 145
QY 257 -----GEEKQKEFTKSPSPSPVNSSETTSFKKFFTHGWAG-----WRKKTSPKK 303
Db 146 EERVGLNQAACAPRLPAPRPAPAPAPEVEELARRLGEAWRGAVRGYQERVAHMETSLDQ 205
QY 304 SKEDDLTAEKKE--QEAKEVDEBEKEKTEPASEBOEPAEDTDQARLSADYKVKELPLE 361
Db 206 TRERLARAVQGEAREVRELELQQLQAERGGLLERRAALQERLEGRWQERLAT-EKFQLAWE 264
QY 362 DQVGDLSEASSEKCAPLATEVDFDEKMEAHQEVVAEVHVST-----VEKTEBEEQGGG-- 413
Db 265 -----ALEQKEQGLSQIOAVLEGRQOLAHKWSLSLEVATYRTLLEAENSRLQTPGGGSK 320
QY 414 -----EAGGVVVVEG--TGESLP---PEKLAEPQEVQAEAB-PAEELMKSRMCMV 457
Db 321 TSLSFQDPKLELQFPPTPGRLGSLLPVLSPTSL--PSPLPATLETVPYPAFLKNGEFLQ 378
QY 458 SGGDHNTQLTDLSPBEKTLPKHPEGIVSEVEMLSQERIKVQGSPLKLFSSSGLKLSGK 517
Db 379 ARTPTLASTPIPP---TPQAPSPAV-----DAEIRAQDAPLSLLQTOGGRQKQAPER 426
QY 518 KQKGRGG-----GGDEEP-GEYQHIHT-ESPESADEOKGESSASSPEPETTCLEK 569
Db 427 LRAREARVAIPASVLPGEPEPGQROEASTGQSPED-----HASLAPPLSPDHS----- 475
QY 570 PLEAPQDGEABEGTT-----SDGEKKREGITP-----WASFCKMVTPKKRVRRPSE 615

Db 476 -LEA-KGSESGSVFSTCRGEGSQIWGLVKEKATIGKVVSVSTQCIWMEEDLNKEI 533
Qy 616 SDKEEELKXVATLSTDSYVSEMDEKTVGE-----EQKEEPKRVDTSVSMAL 669
Db 534 QDSQVPLEK---ETLKSIG---EETQESLKTLENQSHETLERENQCPRLSEEDL--ETL 585
Qy 670 ICVGSSKKRARKASSDDEGGR-----TLGDSHAEERASDKKA----- 710
Db 586 KSLKKNKRAIKGCGSEFSTRKRGROLKPTGKEDTQTLQIKENQBLKMSLEGNLETF 645
Qy 711 ---GTD--AVAPTOEQDQ---AOGSSPEPAGSPSEBGVS----- 744
Db 646 LPPGTNOELVSSLOENIESLTALKENOEPRLSRFEVGDDEKRLPLTENQEPRLSLDE 705
Qy 745 TWSEFKRLVTPRKSKSKLKEKADSVEQLSTE-----LEPSEBSWSIKKIFGR 797
Db 706 NKEAFRLSEKNOBLKTLBEE--DQSIVRPLETENHKSLSLEODQETLRLER--ETQ 762
Qy 798 RKRKADGQEOATVBDSGFVEINEDPNVPAVPLSEINAVREKMEAGNTLPLQLLGA 857
Db 763 QRRSLGRODQWTLRP-----PEKYDLEPLKSLDOE----- 793
Qy 868 VYVSEELKTLVHTVSAVIDGTRAVTSVEERSPMISASVTEPLEHTAGAMPVEEVT 917
Db 794 -----TARPLENENOEPKLSIKES 813
Qy 918 EKDIETAEPTVLTQTLPEKGAHDMDWTSVDFTSEAVATETSEALRTVEETASGAE 977
Db 814 VEAVSLETFEILSKSAGQ-----NLETLSPTQAPLWPEELNINSQNE 861
Qy 978 TTDWVSASQTLDSPTTEATPVOVESGVID-----TEER 1014
Db 862 SSRKNSTRITGVCGSEPRDIQTPRG-ESGILIEISGMEPGEFELSRGVDKESORNLBER 920
Qy 1015 EMOQTALIQAVADKYKEEQ--VPAQTVOYRTSKALEXEVEVEEDEV--LASEKADVWP 1072
Db 921 ENVGGEYVESIRSLSEBQOELPQSHADVOR--MEDVYKQOELAOESPPMAGVENKD--- 976
Qy 1073 KQPVQACAEHLAOGSETQATPESLVEPEVTADV-----HVATCQVYKLOIMEQA 1125
Db 977 -----EALINLEQOGFTK--EEVVEGELNATEEVPFGESHPENPKRGVLEGA 1029
Qy 1126 VAPESSETLTSETNGS-----TPLADSTADGTOQ---DETIDSQS 1165
Db 1030 SVKGAEGLQDBEGSQOQVTPGLQAPQGLPEALIBLVBDVAPGDDQASPEVMLSEBA 1089
Qy 1166 KATAV-----ROSQVTEEAATQKEPSTLPNNVPAOE----- 1200
Db 1090 MGESAAGAPGLGQGVGLGDPGHLTREEMERPLEESLEAKRVOGLHGRKDLLEAGG 1149
Qy 1201 ---EGEGBG--RDVLEPTQO---ELTA---AAVFLAKTEVQEGEVMDLGEKV 1245
Db 1150 LGTERSELFGKSRDPEPRRGRESEAPARAEBAFAETLGHGSGAPSPWPLGSE 1209
Qy 1246 KEEO-----EVEVHSGPNSQKADVITYDSEVNGVAGCOKESTEVQSLSEBG 1293
Db 1210 AEDDVAPVLVSPSTYTPILLEDAFGLQOAEBSQFAS--WGVQRAEAGKVESQOELGSG 1268
Qy 1294 EMETDVKEKREKRPQVVEEGQOETAPENH--GTYGK-----PVLITDMPSEBG 1343
Db 1269 EIPGELQEGEBESRE--SBEDELGETLPDSTPLGYLSPSPSPMTPLESRGHPKTEG 1326
Qy 1344 K-----ALGSLG--GSPSLPDQKAG--C-----LEVQVOSIDTYVT---OTAEA 1381
Db 1327 KEGWDPVALVASEGLEPSEKEGEBEGEGGRSDLSSEEDLDGTAPLPGVGVARE 1386
Qy 1382 VEKVLJETVV-----ISETGESE-----CVGAHLPL 1407
Db 1387 LQGVPLLLDPAAMDRDGEDGPADEBSGEGEEDQOEBRGEGAGRWGGSVGS--LQ 1444
Qy 1408 AAKSSATGHWITQHAEDVYPLGPESQABSIPIYVPAPESTLHPDQ--GELISAQORR 1465

Db 1445 ALSSSGREBLESDSVSVPMW-----DSLRGAVAGAPKTALETESQDASBPSSGEES 1499
Qy 1466 S-----EEDK-----PDAGPDADKESATLEKVLKAEPELLEKSKNKI---V 1507
Db 1500 DVSLEBEDKPGPLEIPSGMEDAGPAD-----ITGVNGQGNLKGSKSHVNGCV 1550
Qy 1508 LNVITAVDQFARTETAPE 1526
Db 1551 MNGLEQSESGARNALVSE 1569

RESULT 15
US-09-864-761-36182
Sequence 36182, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aeomica-X-1
CURRENT FILING DATE: 2001-05-23
PRIOR FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 36182
LENGTH: 617
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC005529.7
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 0.94
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1

; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.6
; OTHER INFORMATION: SWISSPROT HIT: P12036, EVALUUE 2.00e-33
; OTHER INFORMATION: EST_HUMAN HIT: BF445754.1, EVALUUE 3.00e-22
US-09-864-761-36182

Query Match		4.4%	Score 351.5;	DB 10;	Length 617;
Best Local Similarity		23.4%	Pred. No. 1e-08;		
Matches 175;		Conservative 107;	Mismatches 258;	Indels 207;	Gaps 30;
QY	28	GHGPAEASGAAGPADADPATKLPOKNGQLSSVNGVAEQDHHVQENOBQGEVEVVDE	87		
DB	11	GFGP-IPFSLPEGLPKIPSVSTHIKVKSEKIKVVEKSEKETVIVEEQTEQ-----VTE	65		
QY	88	DVGQRESEDEVREKDRVEEMANSTAVEDITKGOEETSEIIEIQIPASENNVEEMVQPAES	147		
DB	66	EVTEEEKEAKEEKEEGEGE-----BEEAEGGEETKS-----PPAEE	105		
QY	148	QANDVGFKVFKFVGFKFTVKDKNEKSDTVQLLTIVKKBEGEGAEASVGAGDHQEPSVET	207		
DB	106	AAS-----PEKEAKS-----PVKEEAKSPAEA-----KSPEKEE	134		
QY	208	AVGESASKESELKOSTEKQETGLKQESSTE-----IPLQAESDQAAEEBAKDEGEKQE	262		
DB	135	AKSPAENVKSPAKAKSPAKEEAKSPPEAKSPKEEAKSPAENVKSPPEAKSPAKEEAKSPA	194		
QY	263	-----KEPTKSPSPSPVNSSETTSFFKKFTTHGWAGWRKTSFKSKEDDLET	311		
DB	195	AKSPEAKSPVKEEAKSPAEAKSPVKEEAKSP-----AEVKSPEAKSPPTKEEAKS	245		
QY	312	AEKKEQEAENVDEEKEKTEPASEEQEPAEDTDQARLSA-DYEKVELPLEDQVGDLEAS	370		
DB	246	PEKAKSPEKEEAKSPPEAKSPVKAEAESPEKAKSPVKAEAESPEKAKSPVKE-----EAK	300		
QY	371	SEKCC-APLATEVDEPKMEAHQEVVAEVHVSTVETKEEEOGGGEAGGVVVEGTGSLP	429		
DB	301	SPEKAKSPV-----KEEAKSPEKAKSPVKEEAKTPE-----KAKSPV-----KEEAKS	343		
QY	430	PEKLAEPQEV-----PQEAEE-PAEELMKSRMCMVSGGDHT	463		
DB	344	PEKAKSPEKAKTLDVKSPEAKTPAKEEARGPADKFPPEKAKSPVKEEVKSPEKAKS----	399		
QY	464	QLTDLSPPEKTLPKHPGIVSEVEMLSQERIKVQGSPLKLPSSSGLKKLSGKKQKGR	523		
DB	400	LKADAKAPEKEIIPKKEE-VKSPVKEEKPQEVVKSPKK-----AEEEKAPATPKTEEK	454		
QY	524	GGGDEEPGEYQIHTSPESADQKGESASSPEE-----PEETTCLEKGP-----	570		
DB	455	DSKKEAPK-----EAPKPVKEEKEPAVEKPEKSKVEAKKEEADKKVPTPEKAP	508		
QY	571	--LEAPQDGEABGT-----TSDGEKKREGITPWASFKKMVTPKKRVRRPSESDEELEE	623		
DB	509	AKVEVKEDAKPKTEVAKKEPDADAKAPEKSP--AEKKEAAPKPKDKEEAKKPEKP	566		
QY	624	KVKSATLSSTDSITVSEMQDEVKTVEEQKPEEPKRRVDTSVSWEALICVSSKKRKARKAS	683		
DB	567	K-----TEAKAKEDDKTILSKC--PSKPK-----AEKAEKSS	595		
QY	684	SSDDEGGPRTLGGDSHRAEASDKDEA	710		
DB	596	STQK-----DSKPEKATEDKAA	614		

Search completed: December 13, 2002, 00:28:36
Job time : 248 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 13, 2002, 00:06:25 ; Search time 59 Seconds
(without alignments)
2600.518 Million cell updates/sec

Title: US-09-902-432-4
Perfect score: 8073
Sequence: 1 MGAGSSTQSPQAGSDT.....AWAQRKCLRLQLKAPVSK 1596

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 73:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	ID	Description
1	5421.5	67.2	1346	2 A57376	probable regulator
2	4083.5	50.6	1684	2 JW0057	gravin - human
3	585.5	7.3	1829	2 T24583	hypothetical prote
4	574	7.1	5327	2 T13564	microtubule-associ
5	501	6.2	2464	1 QRMSP1	microtubule-associ
6	497.5	6.2	3488	2 T34418	hypothetical prote
7	493	6.1	3924	2 S37431	ankyrin 2, neurona
8	489	6.1	6642	2 T29757	protein UNC-89 - C
9	488	6.0	5170	2 T15348	hypothetical prote
10	485	6.0	2484	2 T26216	hypothetical prote
11	485	6.0	2607	2 T26215	gene 11-1 protein
12	483.5	6.0	1948	2 S00485	hypothetical prote
13	470	5.8	2364	2 A56577	microtubule-associ
14	466.5	5.8	1871	2 D96796	probable heat choc
15	456.5	5.7	1804	2 T34518	nestin - golden ha
16	438	5.4	1128	2 T30296	R27-2 protein - Tr
17	434	5.4	1526	2 A45605	mature-parasite-in
18	433.5	5.4	2774	2 A43359	microtubule-associ
19	425	5.3	1110	2 T51116	NF-180 - sea lamp
20	419	5.2	3187	2 JC5837	364K Golgi complex
21	411.5	5.1	3582	2 A47171	chondroitin sulfat
22	408.5	5.1	4687	1 A39638	plectin - rat
23	407.5	5.0	2139	2 T18296	myosin heavy chain
24	403.5	5.0	1909	2 A45592	liver stage antige
25	402.5	5.0	7962	2 I38346	elastic titin - hu
26	402	5.0	2218	2 B84683	hypothetical prote
27	401.5	5.0	4574	2 G02520	plectin - human
28	400.5	5.0	4684	2 A59404	plectin [imported]
29	400	5.0	1830	2 A37981	microtubule-associ

30	400	5.0	3225	2 I52300	giantin - human
31	397	4.9	3507	2 T34513	hypothetical prote
32	395.5	4.9	1825	2 T31507	microtubule-associ
33	393	4.9	3259	1 A56539	giantin - human
34	392.5	4.9	1822	2 S33441	EF protein - Strept
35	389	4.8	1262	2 T2523	hypothetical prote
36	389	4.8	1621	2 A82255	hypothetical prote
37	388	4.8	1805	2 A34736	nestin - rat
38	385.5	4.8	1087	1 QPM5H	neurofilament trip
39	382.5	4.7	1320	2 JC5630	TCOF1 protein - mo
40	382	4.7	2116	2 A26655	myosin heavy chain
41	380	4.7	1558	2 B71603	RESA-H3 antigen pF
42	379	4.7	2094	2 S33124	tpr protein - huma
43	379	4.7	2722	2 T20532	hypothetical prote
44	379	4.7	2738	2 E88320	protein F0A11.6
45	377	4.7	1634	2 T26517	hypothetical prote

ALIGNMENTS

RESULT 1

A57376

probable regulatory protein 322 - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 08-Feb-1996 #sequence_revision 08-Feb-1996 #text_change 08-Feb-1996

C:Accession: A57376

R:Lin. X.: Nelson, P.J.; Frankfort, B.; Tomblar, E.; Johnson, R.; Gelman, I.H.

Mol. Cell. Biol. 15, 2754-2762, 1995

A:Title: Isolation and characterization of a novel mitogenic regulatory gene, 322, which

A:Reference number: A57376; MUID:95257957; PMID:7739556

A:Accession: A57376

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-1346 <LIN>

A:Cross-references: GB:U23146

C:Genetics:

A:Gene: 322

Query Match 97.2%; Score 5421.5; DB 2; Length 1346;

Best Local Similarity 91.6%; Pred. No 5.5e-187;

Matches 1098; Conservative 24; Mismatches 62; Indels 15; Gaps 7;

QY	387	MEAHQEVAVHVS	TEVEKEE	EQGGGEA	GGVVG	TGTE	GESLP	PEKLA	EPVQ	BAEPA	446
DB	1	MEAHQEVAVHVS	TEVEKEE	EQGGGEA	GGVVG	TGTE	GESLP	PEKLA	EPVQ	BAEPA	60
QY	447	EELMKREMCVSG	GDHTQ	LDLS	PEKTL	PKH	PEGIV	SEVEM	LSQER	IKVQ	SPLKLF 506
DB	61	EELMKREMCVSG	GDHTQ	LDLS	PEKTL	PKH	PEGIV	SEVEM	LSQER	IKVQ	SPLKLF 120
QY	507	SSSGLKLSGKK	QKRG	GGDE	EPGEY	QHI	HTES	PESAD	EQKGS	ASSP	PEETTC 566
DB	121	SSSGLKLSGKK	QKRG	GGDE	EPGEY	QHI	HTES	PESAD	EQKGS	ASSP	PEETTC 180
QY	567	EKGPLEAPQDGE	AEETG	TTSD	GKKRE	---	GITP	WAF	KKWV	TKRV	RRPSESDEEEL 622
DB	181	EKGPLEAPQDGE	AEETG	TTSD	GKKRE	---	GITP	WAF	KKWV	TKRV	RRPSESDEEEL 234
QY	623	EKVKSATLS	TDSTV	SEM	QDEVT	VGEE	QKPE	PKRV	DTVS	WEAL	ICVGSKKRKARKA 682
DB	235	EKVKSATLS	TDSTV	SEM	QDEVT	VGEE	QKPE	PKRV	DTVS	WEAL	ICVGSKKRKARKA 294
QY	683	SSSDDEGGP	PTLGD	SHRA	EAS	KDKE	AGTD	AVP	ASTQ	EOAQ	GSSSPAGSPSSEGE 742
DB	295	SSSDIR	-GPTL	GGOS	RG	QOR	SR	TD	AVP	ASTQ	EOAQGSSPPAGSPSSEGE 353
QY	743	VSTWSEFKRL	VT	PRKKS	KS	LEEK	-AED	SSVE	QSTL	IEP	SREESWVS--IKKPIPGRRK 799
DB	354	VSTWSEFKRL	VT	PRKKS	KS	LEEK	-AED	SSVE	QSTL	IEP	SREESWVS--IKKPIPGRRK 413
QY	800	KRADGQEQAT	VS	EDS	GP	VEIN	DDPN	VAV	PLSEY	NAVER	KMEAGQNTLFPQLGAVY 859

Db 414 KGQMRQEQATVEDSGPVEINDEPDVPAVNVPLSEYDAVERKMAQGNALPSCMGCV- 472
Qy 860 VSEELSKTLVHTVSVAVI DGTAVTSVEERSPSWISASYTEBLEHTAGAMPVEEVTEK 919
Db 473 VSEELSKTLVHTVSVAVI DGTAVTSVEERSPSWISASYTEBLEHTAGAMPVEEVTEK 532
Qy 920 DIIAETPVLTOTLEPGKADHDMTSEVDFPTESEAVTATETSEALRTEVETASQAEEET 979
Db 533 DIIAETPVLTOTLEPGKADHDMTSEVDFPTESEAVTATETSEALRTEVETASQAEEET 592
Qy 960 DMVSAVSQUTLSDPTTEATPVQEVESGVLDTEEEERQOAILQAVADVKESEVPAQO 1039
Db 593 DMVSAVSQUTLSDPTTEATPVQEVESGVLDTEEEERQOAILQAVADVKESEVPAQO 652
Qy 1040 TVQRTGSKALEKVEVEEDSEVLASEKEKDVMPKGPVQAGAEHLAQSGSETQOATPESLE 1099
Db 653 TVQRTGSKALEKVEVEEDSEVLASEKEKDVMPKGPVQAGAEHLAQSGSETQOATPESLE 712
Qy 1100 VPEVTADVHVATCOVYIKLQQLMEQAVAPESSETLTDSETNSTPLADSDTADGTQOEBT 1159
Db 713 VPEVTADVHVATCOVYIKLQQLMEQAVAPESSETLTDSETNSTPLADSDTADGTQOEBT 772
Qy 1160 IDSQSKATAAROSQVNEEBAATAQKEPSTLPNNVPAQEEHGEPRGVDLEPTQOEBT 1219
Db 773 IDSQSKATAAROSQVNEEBAATAQKEPSTLPNNVPAQEEHGEPRGVDLEPTQOEBT 832
Qy 1220 AAAPVLAKEVEGOBGEVMDLDEKVEKEQEVFVHSGPNSQAKADVTYDSEVAGVACQGE 1279
Db 833 AAAPVLAKEVEGOBGEVMDLDEKVEKEQEVFVHSGPNSQAKADVTYDSEVAGVACQGE 892
Qy 1280 KESTEVQISLEBGEVETVEKEKRETKPEQVSEBGEQTAAPENHGTYGKVLTLDMPS 1339
Db 893 KESTEVQISLEBGEVETVEKEKRETKPEQVSEBGEQTAAPENHGTYGKVLTLDMPS 952
Qy 1340 SERGKALGSLGSPSLPDDQKACIEVQVQSLDTTYTQAEVAKETIEVTVSETGESPE 1399
Db 953 SERGKALGSLGSPSLPDDQKACIEVQVQSLDTTYTQAEVAKETIEVTVSETGESPE 1012
Qy 1400 CVGAHLLEPKSSATGHTLQHAEDTVPLGPESQAESIPITVPAESTLHPDLOGEIS 1459
Db 1013 CVGAHLLEPKSSATGHTLQHAEDTVPLGPESQAESIPITVPAESTLHPDLOGEIS 1072
Qy 1460 ASQRESEEDKPDGPDADGKESTALEVYLKAEPELLELESKNTVUNVITQAVDQPA 1519
Db 1073 ASQRESEEDKPDGPDADGKESTALEVYLKAEPELLELESKNTVUNVITQAVDQPA 1132
Qy 1520 RRETAPEHTAYVSOVQVPCRLDSREPNNCWTKMDAKMKHPVOPREDLOVLTIVLEAM 1578
Db 1133 RRETAPEHTAYVSOVQVPCRLDSREPNNCWTKMDAKMKHPVOPREDLOVLTIVLEAM 1191

RESULT 2
JM0057
Gravlin - human
C:Species: Homo sapiens (man)
C:Date: 13-Jun-1998 #sequence_revision 10-Jul-1998 #text_change 08-Oct-1999
C:Accession: JM0057
R:Saco, N.; Kokame, K.; Shimokado, K.; Kato, H.; Miyata, T.
J: Biochem. 123, 1119-1126, 1998
A:Title: Changes of gene expression by lysophosphatidylcholine in vascular endothelial C
A:Reference number: JM0057; MUID:98269042; PMID:9604001
A:Accession: JM0057
A:Molecule type: mRNA
A:Residues: 1-1684 <SAT>
A:Cross-references: DDBJ:AB003476; NID:G2081606; PIDN:BAAL19927.1; PID:dl020716; PID:G208
C:Comment: This protein regulates cell growth.
F:433-439/Region: nuclear location signal
F:522-527/Region: nuclear location signal
F:591-596/Region: nuclear location signal
F:671-676/Region: nuclear location signal

Query Match 50.6%; Score 4083.5; DB 2; Length 1684;
Best Local Similarity 58.3%; Pred. No. 5.6e-139;

Matches 903; Conservative 179; Mismatches 378; Indels 89; Gaps 30;
Qy 89 VQGRSEEDVREKRDVEEMANSTAVEDITTKQGEETSEIILQIPASNNVEMVQPAESQ 148
Db 9 VQGRSEEDVRSKSDDKEMATKSAVAVHDITDQGEETSEIILQIPSSSSNILELTQPTESQ 68
Qy 149 ANDVGFKKVFEPKFPVKKDKNEKSPVQVLTQVKKDEGGAGASVAGAGHQSPSVETA 208
Db 69 ANDVGFKKVFEPKFPVKKDKNEKSPVQVLTQVKKDEGGAG--AGAGHKDPSL--G 123
Qy 209 VGESASKESELTQSTKQTEKQEGTLKQESSTELPLQAESDQAEEBAKDEGEKQEKPTKS 268
Db 124 AGEPAKSESEKQSTKQTEKQEGTLKQESSTELPLQAESDQA--EBCKEGEKQEKPEKS 182
Qy 269 PESPSVNSSETSSFKKFTTHGAGWRKKTSPKSKEDDLETAKEKKEQAEKVDEEEK 328
Db 183 AESPTSPVTSSETSSFKKFTQGWAGWRKKTSPKXEDVEASEKKEQEKPEKVDTEED 242
Qy 329 EKTPEASE-----QEPADTDQARLSADYEKVELPLEDQVGLDEASEKCAPLA 379
Db 243 GKAEVASEKLTASDAQHPQEPASEABEPRLSABYEKVELPEEEOVSGSQGEPEKRAPLA 302
Qy 380 TEVPDEKKEAHQ-EVVAEVAHVTVEKTEEPQGGGGEAGGVVVGTSGLPPEKLAEPQE 438
Db 303 TEVPDEKTEVHQEVEVVAEVAHVTVEBERTEQ-----KTEVETAGSVABELVEMDA 354
Qy 439 VPQEAPEELIMKSRMCMVSGDHTQTLTDSPEEKTLPKHEBGIVSEVEMLSQERIKVQ 498
Db 355 EPQEAPEAKEVLKTEKTCVSGEDPTQAGADLSPDEKVLSPREGVAVSEVEMLSQGRMKVQ 414
Qy 499 GSPKLKLFSSGLKLKSKKKQKRGCGGDEPEGEYQIHTHESPESADQEGESASAPPE 558
Db 355 EPQEAPEAKEVLKTEKTCVSGEDPTQAGADLSPDEKVLSPREGVAVSEVEMLSQGRMKVQ 414
Qy 415 GSPKLKLFSTGLKKLKSKKKQKRR-GGGDESEGHQVPAVPSQPSQEGQSSASSAPPE 473
Db 559 EPEETTCLEKPLLEAPQGEAEEGTSSDGKKRREGITPMASFKKMWVTPKKKVRPSESDX 618
Qy 474 EPEETTCLEKPLLEAPQGEAEEGTSSDGKKRREGITPMASFKKMWVTPKKKVRPSESDX 533
Db 619 EEELEKVKSAITLSTDSVSEMDEKTVGEEQKPEEKKRRVDTSVSWEALICVSSSKR 678
Qy 534 EDELKVKSAITLSTDSVSEMDEKTVGEEQKPEEKKRRVDTSVSWEALICVSSSKR 593
Qy 679 ARKASSDDEGPRTLGSDSHRABASDKACAGTDAVAVASQEDQAGSSSPAPAGPS 738
Db 594 ARKASSDDEGPRTLGSDSHRABASDKACAGTDAVAVASQEDQAGSSSPAPAGPS 653
Qy 739 EGEGVSTSEFPRLVTPPKKSKSKLEEKAEQ-----SSVEQSTELTEPERSWSVTKKEI 794
Db 654 EGEGVSTSEFPRLVTPPKKSKSKLEEKSEDSIAGSVGEHSTPTEPKEKSEWSVTKKEI 713
Qy 795 PGRKKRADGKQEQATVEDSGPVEINDEPDVPAVNVPLSEYNAVERKME--AQGNTEL 851
Db 714 PGRKKRADGKQEQATVEDSGPVEINDEPDVPAVNVPLSEYDAVERKMEQAQAKSAEQ 773
Qy 852 POLGAVVSEELSKTLVHTVSVAVI DGTAVTSVEERSPSWISASYTEBLEHTAGAMP 911
Db 774 POLGAVVSEELSKTLVHTVSVAVI DGTAVTSVEERSPSWISASYTEBLEHTAGAMP 833
Qy 912 PVEETEKDIIA-EETPVLTOTLEPGKADHDMTSEVDFPTESEAVTATETSEALRTEEVT 970
Db 834 LTEETLESEVLAEEPPVTEPLPENRKARGDVYSEALTEPEVNTAETAGPLCAEAGT 893
Qy 971 EASGAETTDVMSAVSQTLDSPDTEATPVQEVESGVLDTEEEERQOAILQAVADVVK 1030
Db 894 EASGAETTDVMSAVSQTLDSPDTEATPVQEVESGVLDTEEEERQOAILQAVADVVK 953
Qy 1031 EESQVPAT-----QVQRTGSKALEKVEVEEDSEVLASEKEKDVMPKGPVQAGAEH 1083
Db 954 EESQVPAT-----QVQRTGSKALEKVEVEEDSEVLASEKEKDVMPKGPVQAGAEH 1007
Qy 1084 LAQGETGOATPESLE-VPEVTADVH---VATCOV-----IKLQQL-MEQAVAPESSET 1133
Db 1008 FTQGVAVGQTTPESEKAPQVTESIESSELVTTQAEITLAVKYSQEMEQAIIPDSVET 1067

Qy	1134	LTDSETNGSTPLADSDTAGTQDQDETI	DSQDSKATAAVRSQVSTEEBAATAQKEB	ESTLP	1193				
Db	1068	PTDSETDSTPVADFDAPGTTQKDEIVE	THEENEVASGTQSGGTEAEVPAQKERPPA	-P	1126				
Qy	1194	NNVPAQEEHGREPG-RDVLEPTQOELTAA	VPVLAKTEVQOEGEVDWLGEKVK	-----E	1247				
Db	1127	SSFVQOETKEQSKMEDTLLEHTDKESV	ETVTSLLSKTEGTQ--EADQYADEKTKD	VPFFE	1184				
Qy	1248	EQEVFVHSG--PNSQKAADVTDSEVMG	VAGQOKEKSTEVQS-----LSLEE	GEMETDVE	1300				
Db	1185	GLEGSIDTGITVRSREKVTAEALKGEGE	TEAECKDDALEQSHAKSPSPVEREMVQVE		1244				
Qy	1301	KEKRETKPEQVSEEG-EQETAPEHEGTT	GKPVLTLDMSERSEKALGSLGGS-PSLPDQ		1358				
Db	1245	REKTEAEPTHNVEEKLHEHTAVTVSEV	SKQLLTQVNPVIDGAKEVSSLEGSPPCLGQ		1304				
Qy	1359	DKAGCIEVQVQSLDVTVTOTAAAEVKI-	ETVWISGETSPCEVCAGAILLPARKS	SATGG	1416				
Db	1305	EEAVCTKIQVQSSEASFLLTAAAEKVL	GETANILLETLETPAGAHVLEKSEKNE		1364				
Qy	1417	HWTLOHAEDTVPLGPESQAESPIIVTP	APESLTHPDQCEISASORSEEBDKPDAGP		1476				
Db	1365	DFAAHPGEDAVPTGPDCAKSTPVI	SATTKXGLSSDLEGEKTTSLKWKSD	EVDEQVAC	1424				
Qy	1477	DADGKESTAIEKVLKAEPE--LLELES	KNKIVLNVIOAVDOPART-ETAPETH	AYDSQ	1533				
Db	1425	EV-KVSVAIEDL--EPENGILELETKS	KLQVNIQTAVDQFVTEETATMLTSELQ		1479				
Qy	1534	TQVPACRLDSREPNRCWK-----	MKDAKMKHPVQOPREDQLVLTLEA		1577				
Db	1480	TQAHVIKADSQDAGETEKEGEBPLAS	AODEPTITSAKESESTAVGQA		1528				
RESULT	3								
T24583									
hypothetical protein T06D8.1 - <i>Caenorhabditis elegans</i>									
C:Species: <i>Caenorhabditis elegans</i>									
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999									
C:Accession: T24583									
R:Palmer, S.									
submitted to the EMBL Data Library, April 1995									
A:Reference number: Z19909									
A:Accession: T24583									
A:Status: preliminary; translated from GB/EMBL/DBJ									
A:Molecule type: DNA									
A:Residues: 1-1829 <WIL>									
A:Cross-references: EMBL:Z49130; PIDN:CAA88964.1; GSPDB:GN00020; CESP:T06D8.1									
A:Experimental source: clone T06D8									
C:Genetics:									
A:Gene: CESP:T06D8.1									
A:Map position: 2									
A:Introns: 1391/3; 1432/3; 1470/3; 1505/1; 1520/1; 1616/1; 1644/1; 1687/3; 1742/1									
Query Match 7.3%; Score 585.5; DB 2; Length 1829;									
Best Local Similarity 23.2%; Pred. No. 1e-13;									
Matches 375; Conservative 174; Mismatches 662; Indels 405; Gaps 65;									
Qy	2	GAGSSTE-----QRSPEQ	PAGSDTPSELVLSGHGPAAEASG-----AAGDPADADPAT--	49					
Db	106	GSGETTVVAVVSSGEEPPASSSTS	VPTELSKDDQVTEASGETTTTAAATEASSEETTTTS	165					
Qy	50	KLPQKNGQLSSVNGVAE-----	QGDVHVQENQEGQEEF-----VVD	EDVQ	91				
Db	166	AVTEGSGGETTTTSVTEASSEAT	TTTTPAGTEASGEETTTTSAVTEGSGEETTVVAVVSSGE	225					
Qy	92	RESE-----DVREKDR	VEEMAANSSTAVEDITKQOEBETSEIIQIPASENNVEBW--	142					
Db	226	EPASSSTSIPTELSKNDQVTEAS	GEETITAAATEASEETTTSAVTEGSGEDITTVAVVEL	285					
Qy	143	---QPAESQANDVPKQKVPK	FGVGFPTVKKDNKNEKSDTVQLLTJTVKKDGE	GAEASVGAG-	198				
Db	286	SQGEQPASS-----TSIPT	ELSKDDQV-----TEASGEETTTAAAT	321					

QY 1207 GRDVLPEPTQOELTAAAVPLATKEVQGEVDMLOGEKYEBOEVFVHSGSPNSQAKADVT 1266
 Db 1162 G-----EEDDAPFPVTGAPLDPSTTEASVSTSA--ITTDETSVAADBESTSTSAAGEVQ 1231
 QY 1267 YDSEVWGV--CCQEKESTEVQSLSEEGEMETDVEKEKRETPKEQVSEEGEOETAAPEH 1324
 Db 1232 SSSAIIIDSTVASEBQTSSEATSVIESSGE-EVTTTDEMLVTVSTVAQLEGGGITAABEK 1290
 QY 1335 EGTYGKPVLTLMPSSEKALGSLGGSPS--LPDODKAGCIEVVOYQSLD---TTVTQT 1378
 Db 1291 D----EDSVTTEATSGSTTVSSSDSGESTVAPNDSESTTSSSQSTDESGVTAAES 1346
 QY 1379 AAVEVLE--IVVISEGSESPCEVAHLLPPEKSAATGGHMTLQAAEDTV---PLGPES 1433
 Db 1347 KQESSSTTAPAPVTVSTKISGSE-----DEEDSPTEHFTIGIDETMFKSLVPT 1397
 QY 1434 QAESIPITTPAPESTLHPDLQGEISASQERSEEDKPDAGDADGKSTAIKVV 1489
 Db 1398 HREDLPNVGVFPSPSEPKKNPDE-----EEHEEDDGTKSDYEDNVSKTI 1444

RESULT 4

T13564
 microtubule-associated protein homolog - fruit fly (Drosophila melanogaster)
 N:Alternate names: hypothetical protein EG:49E4.1
 C:Species: Drosophila melanogaster
 C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 17-Nov-2000
 C:Accession: T13564
 R:Spanos, L.; Papadogiannakis, G.; Siden-Kiamos, I.; Louis, C.
 Submitted to the EMBL Data Library, April 1999
 A:Description: Sequencing the distal X chromosome of Drosophila melanogaster.
 A:Reference number: Z17689
 A:Accession: T13564
 A:Stature: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-5327 <SPA>
 A:Cross-references: EMBL:AL031128; PIDN:CAA20006.1
 C:Genetics:
 A:Cross-references: FlyBase:FBgn0025392
 A:Introns: 24/2; 52/3; 104/3; 179/1; 232/1; 1669/3; 2566/1; 4798/3; 5272/1
 A:Note: EG:49E4.1
 C:Superfamily: Drosophila 576k microtubule-associated protein homolog

Query Match 7.1%; Score 574; DB 2; Length 5327;
 Best Local Similarity 21.6%; Pred. No. 9.2e-13;
 Matches 415; Conservative 285; Mismatches 734; Indels 484; Gaps 82;

QY 6 STPEQRPQPDGSDTSELVLSGHGPAAEASGAADPAD-----ADPATKLPKNGQL 58
 Db 2744 SYVESKDKDAEKSESRESPIASGEFVPRESKSLPDSKDTSRGVSVEYTADEKSEBOQ 2803
 QY 59 SAVNGVAE--QGDVHVOENQOGEVEVDEVDGQRESEDPREKRDVEEVA---ANSTA 112
 Db 2804 SRRESVAESVKADTKDKGSGEASRPSVDL---KDDDEKESRQSTTSGHKAMSTM 2860
 QY 113 VEDITDQGEETSEITIEQIPASNNVEEMVQ-----PAESQANDV--- 152
 Db 2861 GDESPMDKADSKKE-----PSRPSVAESIKHENTKDEESPLGSRDVSVAESIKSDITKG 2915
 QY 153 -----GFKYVFKVGVFKYTKDKDKNSDPTVOLLYTKQDEGAEASVAGDQGE--- 203
 Db 2916 EKSPPLPSKEVSPRESVVGSLKDEKASRRRESVAESVYPESKDATAPPKESKRESVUL 2975
 QY 204 -----SVETAVGESASESEIKOSTE-----KQGTGLKQEOSSTEIPL 241
 Db 2976 GSLKDGDKTTSKRVAVADIKDEKSLIVSQEASRPESEASLKDAAPSOEHSRRESVA 3035
 QY 242 QAESD-----QAEEBAKDEGEKQEKPTKSPSPSPVNSSETTSFKKFP 288
 Db 3036 ESKVDKSPVASKASRPASVAENAKDADSEKORPESIPQSKAGSIKDEKSP----- 3089
 QY 289 THGMAGWRKTKTSFKSKED-----LETAKEKEQ----- 318

Db 3090 -----LASKDEAKESKESRESRESVAEQPLVSKESRNPASVAESVKDEAKESKEEPLM 3143
 QY 319 -----EAKXDEEEK-----EKTE-PASEQOEPAEDTDQARLADYIKV 356
 Db 3144 SKASRPASVAGSVKDEAKESKESRRRESVAEKSPLPSEKASRPASVAESVKDEADSKSE 3203
 QY 357 ELPLBDQGDLEASGEKCAE--LATEVFDE-----KMAHOEVAEVA----- 398
 Db 3204 ESRRESGAKESPPLASKESRASPASVAESIKDEAKESKESRRRESVAEKSPLPSEKASRPTS 3263
 QY 399 -----VSTVEKTEEBQGGGAEAGVVEGTGESLPPEKLAEPQEPQEAPEELMKSR 453
 Db 3264 VAKSVKDBAEKSKESRRSDVAEKSPPL--ASKASRPASVAE--SVQDEAKSKSE--ESR 3317
 QY 454 EMCVSGGHTQTLDSPEEKLPPKH-PGGIVSEVEMLSQORIK--VQGSPLKLLPSS-- 508
 Db 3318 RESVA-----EKSPLAYKEARPASVAESIKDEAKESKESRRRESVAEKSPPLASKESR 3372
 QY 509 -SGLKLLSGKKQKRGKGGE--EPGEYGHITESPESADQGEASASSPEEPETTC 565
 Db 3373 TVASVSKDEAKSKESRRSDVAEKSPPLASKESRNPASVAESVKDEAKESKESRESV 3432
 QY 566 LEKGPL---EAPQDEAEEGTTSDEKKREGITPWSFKKMTPKRVRR----- 613
 Db 3433 AEKSPPLASKESRNPASVAESVKDDEAKESKESRRRESVAEKSPPLASKESRNPASVAESVKD 3492
 QY 614 -SESDKEE--ELEKVKATLSTDSVSEMODEVKTVEGQKPEPKRRVDTSVSEAL 669
 Db 3493 EAEKESKESRRRESVAEKSPPLSKESRPTSVASVKDEAKSK--EESRRE--SVAEKSS 3548
 QY 670 ICVSGSKRRARKASSDDEGGPRTLGGDSHRAEASKKDEKGTVA--PASTQE--QDPAQ 726
 Db 3549 IASKASRPASVAESVKDEA--EKSKESRRRESVAEKSPPLASKESRNPASVAESVKDEAL 3606
 QY 727 GS-----SSPEAGSPSEGGVSTWSEFRKLTVPKKSKKLEKAEADSSV--- 772
 Db 3607 KSKEVSRRESVAEKSPPLSKESRPTSVVA--ESVMD--EADSKESRRRESGAKESPLASM 3663
 QY 773 -----EQLSTIEPSREBS--WVSIKFTIPGRKKR---ADGKQEA--TVEDS 814
 Db 3664 EASRPTSVASVKDETEKSKESRRRESVTEKSPPLSKESRPTSVASVKDEAKESKES 3723
 QY 815 GPVELNEDDP-----NVPAYVPLEYNAVVEREKQAQNTLPLDAAVYVSEELSTL 868
 Db 3724 RRESVAEKSPPLASKESSRPASV-----ASIKDEAGTQK-----EERRESM 3765
 QY 869 VHTVSAVAIDGTR-AVTVEERSPSMISASVTEPLEHTAGEAMPVEVTEKDIIA--- 923
 Db 3766 PESGAESIKQDQSSLAKETSRPDSVVESVKDETEKEGSAIDKSQVASRPESVAESAK 3825
 QY 924 -ETTPVLQTLPEGKDAHDMDVTSSEVDTSEAVTATETSEALRTBEVTASGAETTMV 982
 Db 3826 DEKSPPL--HSRPE-----SVADKSPDA--SKESRSLSVAE--YASSPIEERG- 3867
 QY 983 SAVSQTLTSDPTTEEA-----TPVOESGVLDTEEEERQTAIIQNAVADKYESQ 1034
 Db 3866 RSIADLSPLNLTGAKGKPLTSSPIDVAGEDFLEVAAESSPRPAVLSPKPEFSQPTG 3927
 QY 1035 VPATQVORTGSKALEKEVEVEDESE---VLASEKEDVMPKGVQVQAGAEHLAQSET 1090
 Db 3928 HTASTPVE--ASPVLEIEIVDQHTSGVGATGAFAFDL---DLTETKSTTVIKQSET 3983
 QY 1091 G-----QATPESLE-----VPEYADVDHATCQVILQQLMEQAAVABESSETLT 1135
 Db 3984 TLFEITLTSKVESKAVLVSSVQVQVEKVTQTSVQAEITTVTSLQELTK-----KSEQLT 4038
 QY 1136 -----DSETNSTGLADS--DTADGTQODETIDSDQSKATATAVAQSVQTEBEATPAQKE 1187
 Db 4039 EIKSVLDITNISVNTLFTAVETIKKVQDVT--EKVIEKATEHVSHTTTGSESTTSQ 4097
 QY 1188 EPSTL-----PNNVPA-----QEEHGEPEGRDV----- 1210
 Db 4098 EKSSLDLTGTFSLRLETHITTVGSPPEFTVIGCRDPLVLAHDIKEDEBHRFPSPSDVDKAA 4157

QY 1211 -----LEPTQOELTAAAPVLAKEVQGEVDWLDG-----EKVKEQEVFVH 1254
 Db 4158 IIPQPMRPLSPREEV-AKIVADVAKVLKSDKDITDIIPDFBERQLKEELKSTADTEEE 4216
 QY 1255 SGPNs--QKAADVTYDSEVMGVAGCQEKSTEVQS--LSLEEGEMETDVEKEK-----R 1304
 Db 4217 SDRSTRKSLKSVKVEIE-----SEKSSPDQKSGPISIEEKDKIQSEKAQLRQIGILA 4271
 QY 1305 ETPEQVSEGE-----QTAPEHEGTYGKPVLLTDMPSRGRKALGSLGGSPSLPDQ 1358
 Db 4272 SSRPESVASQSPSPSQAASHEHK-----EVELSESHKAESK-----SRPES 4316
 QY 1359 DKAGCIEVOQSLDTTDTTQTA-----EAVEKVIETVWISSETGESPECVGAHL 1405
 Db 4317 -----VASQVSEKMDKTSRPSASTSQSTKEGDEETESLLHSLTITTTETVETQMBEKS 4371
 QY 1406 LPAEKSSATGGHWTLOHAEDTVPGLPSQASPIIIVTAPAPETLHPDLQGEISASORE- 1464
 Db 4372 FESVSTSVTKS--TVLSSQSTVOLREESTESL-----SSSLKVE-----DSRRRES 4416
 QY 1465 -RSEEDKPDAGPDAGDKETAIEKVLKASPEILELESKNKIVLNVIQTAVDQFARTE- 1522
 Db 4417 LSSLLAEKGGIATNTSLKEDTS-ASASQLELLVQSECSSESISVBIQTSIAQKSNKEI 4475
 QY 1523 -TAPETHAYDSQTVQACRLDSREPNCRTKMDAKMKHPVQPPREDLQVLTVLEAWA 1579
 Db 4476 KDARET-----KVTSQFTTTTSATKDDSLKTVABFLATEKIVSAKEAFS 4521

RESULT 5
 QRMSP1
 microtubule-associated protein MAP1B - mouse
 N:Alternate names: microtubule-associated protein MAP1(X); microtubule-associated protein
 C:Species: Mus musculus (house mouse)
 C:Date: 30-Sep-1991 #sequence revision 30-Sep-1991 #text_change 01-Sep-2000
 C:Accession: S07549; S44387; A33645
 R:Noble, M.; Lewis, S.A.; Cowan, N.J.
 J. Cell Biol. 109, 3367-3376, 1989
 A:Title: The microtubule binding domain of microtubule-associated protein MAP1B contains
 A:Reference number: A33645; MUID:90094539; PMID:2480963
 A:Accession: S07549
 A:Molecule type: mRNA
 A:Residues: 1-2464 <NO>
 A:Cross-references: EMBL:X51396; NID:G52999; PIDN:CAA35761.1; PID:G53000
 R:Sanchez, C.; Padilla, R.; Paciucci, R.; Zabala, J.C.; Avila, J.
 Arch. Biochem. Biophys. 310, 428-432, 1994
 A:Title: Binding of heat-shock protein 70 (hsp70) to tubulin.
 A:Reference number: S44387; MUID:94234720; PMID:8179328
 A:Accession: S44387
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 653-663, 'IC' <SAN>
 C:Superfamily: microtubule-associated protein MAP1B
 C:Keywords: microtubule binding; phosphoprotein; tandem repeat
 F:589-786/Domain: microtubule binding #status experimental <MTB>
 F:589-592,639-642,649-652,655-658,660-663,668-671,674-677,679-682,683-686,687-690,691-699
 R:K-E/D-X)
 F:1861-2064/Region: 17-residue repeats
 F:91,116,351,888,1124,1153,1168,1208,1662,1877,1918,2003,2030,2054,2083/Binding site: ph
 F:147,969,1336,1562,1563,1702,1708,1990,2057,2063,2419/Binding site: phosphate (Thr) (cd
 F:1953/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match
 Best Local Similarity 21.1%; Pred. No. 1-5e-10;
 Matches 307; Conservative 266; Mismatches 648; Indels 532; Gaps 86;

QY 45 ADPATKL-----PQKQGLSSVNGVAFQGDVH-----VQEEHQEGEEVWDVGYQRE 93
 Db 478 ANPAEKIIRVFCGNSQVNIQLEKLEKHLDFLQPLATQKLTGQVPTPPVQVKLQK 537
 QY 94 SEDVREKORVEEMAANSTAVEDITKQGEETSEI-----IEQIPASENVEEMVQPAESQ 148

Db 538 RADRES---LKPATXPVASKSVRKESKEETPEVTKTSQVEKTPKVESKE----- 584
 QY 149 ANDVGFKVFKVGFVFTVKDKNEKSDTVQLLVKDKDEGCABASVGAGDHQEPSEVETA 208
 Db 585 -----KVLVKDKPVKTESKPSVT-----EKEVSKKEQSP-VKAE 619
 QY 209 VGESASKSELKOSTEKQSGTLKQOSSTETIPLQASDOAAEEAKDEGEKQEKPTKS 268
 Db 620 VAEQKATESKPKVTKDK---VVKKE-----IKTKLEEKKEKPKKEVVKEDKTPLKK 669
 QY 269 PRSP-SSPVNSETTSFKKFFTHGMWRKKTSTFKSKEDDLTAEKKEQAEKVEDEE 327
 Db 670 DEKPRKEEVKKEIKK-----EERKELKVEVKKETPLKDAKVEVKKEKK---EVK 720
 QY 328 KERTEPASBEQPAEDTQOARISADYKVELPLEDOVGDLSESEKCAPLATEVEFDE-- 385
 Db 721 KEEKPKKEIKKISKDKSTQSDTKPS-ALKPKVAKKEESTKKE--PLAAGLKDKG 777
 QY 386 -----KWEAHQEVVAHVHSTVEKTEEBEOGGGEAGGVVVEGTGESLPPEKLAEPQEV 440
 Db 778 KVKVKKKEGKTTEAATAVGTAAVVAAGIAASGPVKELEAER---SLMSSPEDLT 834
 QY 441 QBAEPAEELMKREMCVSGGDHTQLTDLSPKEK---TLPKHPEGIVSEVEMLSQSERIKV 497
 Db 835 KDFEE---LKAEEIDVAKIKPQLELIEDKELKETQGEAYVIQKETEVSKGSAE--- 887
 QY 498 QGSPLKLFSSGLKSLGKOKGKRGGGDBEPGEYQHIHTESPESADEOKGESSASSP 557
 Db 888 --SP-----DEGITTTEGE-----GECEQTPPELEPVKEQGV--DIEKFEDEGAGF 930
 QY 558 EPEETTCLEKPLEAPQDGEABEGTTSQGEKKREGITPWAFKQMTPKKRVRRSESD 617
 Db 931 EESSET---GDYEEKARTEEAEPEDGEDNASG---SASK-----HSTPDED 972
 QY 618 KEEELB---KVSATLSSTSTVSEMDEKVTGVEEOKPEEPKRVDTSVSWEALICV 672
 Db 973 ESAKAADVHLKESRVSVDRAEDMDVLEKGEAEQSEEGEED----- 1021
 QY 673 GSKKRAKASSSDEBG-QPRTLGGDSHRAEASKDKAGTDAVPASTQEQOAOQSS-- 729
 Db 1022 -----KAEDAREGEYEPDKTEAEDYVMADVAAEAGV-----TEEOGYLGTSAK 1067
 QY 730 -----SPEPAGS-----PSEGEVSTWESPK-----RLVTPRKSKKLEBAED 769
 Db 1068 QFGIOSPSREPASIHDETLPQGSSEATASDEENREDQPEEFTATSGYTQSTIETSESP 1127
 QY 770 SSVQQLST-----EIPSRSEESVSIKKF-----IPGRR 798
 Db 1128 TPDENSTPRDVMSETNNEETESPSQEFVNITKYESSLYSQEYKPAVASFNGLSEGS 1187
 QY 799 KKRADGQKQATVEDSGPVIEINEDDPNPAVVPLSEYNAVEREKMAEQNTLPLQLGAV 858
 Db 1188 TDATDGKYNASASTISPPSSMEEDKFSKA--LRDAYCSEKELKASAELEIKD----- 1240
 QY 859 YVSEELSKTLVHTVSVAVJDGTRAVTSVEERSPSWISASVTEPLEHT-AGE-----AMPP 912
 Db 1241 -VSDE-----RLSPA-----KGPS-LSPPSPSPIEKTPLGERSVNFSLTP 1278
 QY 913 VE-EVTEKDIABEETPVLTOTLPEGKDAHDDMTVEVDFTSEAVTAT-----E 959
 Db 1279 NBIKVSAGEARSVPGVTOAVVEBHCASPEKTLIEWSPSQSVTGSAGHTPYQSPDDE 1338
 QY 960 TSEALRTE-----EVTEASGABETT----- 979
 Db 1339 KSHSLPTEVSENAQVPSVFESEAKDENERASLSPMDEPVPDSSSPVEKVLSPRLSPPL 1398
 QY 980 -----DMVSASVQL-----TDSP-----DTTEATPVQEVES-GVLDTBEEROT 1018
 Db 1399 LGSSEPYEDFLSADSKVLRRESPEFGKNGKQFPDRESPVSDLTSTGLYQDKQEKST 1458
 QY 1019 QAILQAVADKVEE-SQVPATQTVORTGSKALEKVEEVEDSEVLASEKEKDMKPGPVQ 1077
 Db 1459 GFI-----PIKEDFGPEKKTSDVETMSQSALAUDEKRLGGDV--SPTQIDVQSGSPK 1510

QY	1078	EAGNEHLAQSEITQ-ATPESLEPEVT-ADVDHVAIQVILQQLMGOAVALPSSSELT	1133
Db	1511	EDYKMSISEGVSSKSKATPVDEGAEDTYSHNEGASV-----STASVAITSFP	1559
QY	1136	DSEFNGSTPLADSDVADGTOQDEFTIDSQSKATVAAROSQVTEBEAATAOKEE-----	1188
Db	1560	EPTTDVYSPSLHAEV--GSPHSTIEVD--DLSVSVVQGTPTTQETEMSPSKBCEGRPMSI	1615
QY	1189	-----PSYLENNVPAOEHEGEBGRDV-----LEPTOCELTAANPV	1225
Db	1616	SPDPFSSKTAKSRTPVVDHRSBOSMSIEGQSPKHSFAMDPSRQSDPHPTLGS---V	1672
QY	1226	LAKTEVGOEGEVDMLDEKXKEQEOEYFVHSGPMSKAAVYD---SEVMYAGCOEKES	1282
Db	1673	LHITENGPT-EVDY---SPCIDQDSSLHKIPTPEPS-YTQDNDSELISVSQVEAPS	1727
QY	1283	T-----EVOSLSLEGEMETDVEK-EKRE	1305
Db	1728	TSSAHTPSQIASPLQEDTTLSDVVPREPMSLYASLASEKVOISLEGEXLSPKSDISPLTPRE	1787
QY	1306	TKP-----EYVSEGEDETAAPHEGTYGKPV-----LTLD	1336
Db	1788	SSPLYSRPFQSDSTSAKETAA-AHQSSSPPIDATAEYGRSSMLPDTHOHLALNRD	1846
QY	1337	MPS-----ERGKLG-----SLGSPSLPQODRAGCIEYOVOSLDT--TVTQTA	1379
Db	1847	LTTSSVAKDSGKTPGDPNVAQKPEMAASPRBEYD---YEQEKIRHVDVRY	1902
QY	1380	EAVEKVIETVVISSETGESPECVGAHLLPAKKSAT--GGHWTLQHAEDTVLPGRESQAS	1437
Db	1903	EKTERTIKSP--CDSGYSYETI-----EKTTPDEGDGYCEIYEKTRT-PRGGYS	1952
QY	1438	IPILVTAPESTLHPDQGHISASQRESRBE-----EDKPAQGP-ADGKESTAE	1487
Db	1953	YEI-----SEKTRTPEVSG-YTEKTERSRRLLDISNGYDTPEDGHTLDCSYETT	2007
QY	1488	KVLNAEPFLELESKSKKIVLNVIQTAVPDQFARTETAPETHAY-----DSQTPVACRLD	1542
Db	2008	EKITSFP--ESESYS-----YESTKTRTSPDTSAYCVETMEKITKTQASTY	2053
QY	1543	SRE-PNRWCYTKMDAKAKKHPVQPRELQVLTV	1574
Db	2054	SYETSDRCYTTEK---KSP-SEAKQDVDCIV	2081

```

RESULT 6
T34418
hypothetical protein F12F3.3 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C/Accession: T34418
R/Fulton, B.; Wohlmann, P.
submitted to the EMBL Data Library, July 1998
A/Description: The sequence of C. elegans cosmid F12F3.
A/Reference number: Z21521
A/Accession: T34418
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-3488 <FULL>
A/Cross-references: EMBL:U80022; PIDN:AAC5885.1; GSFDB:GN00023; CESP:F12F3.3
A/Experimental source: strain Bristol N2; clone F12F3
C/Genetics:
A/Gene: CESP:F12F3.3
A/Map position: 5
A/Introns: 281/3; 332/1; 562/3; 600/3; 1866/3; 1944/3; 3393/1

Query Match      6.2%  Score 497.5;  DB 2;  Length 3488;
Best Local Similarity 20.1%;  Pred No.3,le-10;
Matches 330;  Conservative 284;  Mismatches 568;  Indels 457;  Gaps 677;

```

Dh	269	RDGLTIDDSGDR--MEVRHED---EMRGMILLIDCKD-EEAEVACQAI-----NVA	316
Qy	139	-----EENVQPAESQANDVGKVFVKFGKFTVKQDN-EXSDTVOLLTVKKDEGEG	190
Dh	317	GEAWCFDVAVVMHSESRD-----DKSVEVDSDTVLEEKQDGD-357	
Qy	191	AEASVGADHOEBSV-ETAVGESASVESTLKSTOKGTQKQ--EGOSTIIPLOASDQ	247
Dh	358	-----DKSKPTKKKI I KKKETESQVTAABEQOKISEVDVDSVAETEYGAKKP	409
Qy	248	AAEE-----EAKDGEERKEKEPTKSPESPSS---PVNSETTSPFKFTMGAMWRKT	299
Dh	410	DAEKPTLSAKKDSKSKSDPEASTEKSTTEKPTNDKS-----KKS	454
Qy	300	SPKSKSDEDLETAERKKEOEBAEKVDEEKEKTEPASEOEPAEDTQARLSADYEXELP	359
Dh	455	AEKTVAPKKEVT--GKPIEAKKPVEDKDKDASQPS-SKESPPTDGKKKQI PKALFIP	511
Qy	360	--LEDQGD-----LEASSER--CAPLATEVDEMEHGVNAEVH	398
Dh	512	DEISSRCFDPSTMHSETNITTTIRGREGSADKPLVPELPSASYMKFTLVESAKEAE	571
Qy	399	VSTYEKTEBEQGGGGEAGGVVVEGTSELPEKXLAEPQ-EVPOEABEAEELMSREMCV	457
Dh	572	FSFKRRSTPDDKSRRKEG-----LPPAKSEKKDEYTAEKQSTELIISKKEV	621
Qy	458	SGCDHTQTLDSPEEKT-----LPRHPEG-----IVSEVENTLSQERIKVQSGPLKLF	506
Dh	622	---DEKSISEQPDKNKKSIVGVPEPKAEPETKVDVSEEV-----PKKTI	667
Qy	507	SSSGIKLISGKKQKGRGGGDEEPEEYOHITHESPESADBEKGSSASPEBEPTCL	566
Dh	668	KKKTEKSDPSISQKSNLTKPADDDKSKSDVDTKSKTTEDQTKVATOSKLEKADTT--	725
Qy	567	EKGPLEAPOESEAREGTTSGEKKRGGIIPMASPFKMVTPKKRVARPESPKEELEKVK	626
Dh	726	KOIELETIVDDKSQKYL--KKKTEKSPFISQSETPPVVEPTKPAESHOKEIAEVNK	782
Qy	627	SATLSTSDTYS-EMODEVTVGEEOKEPEEPKRVDTSVSWEALICVSSSKKARAKASS	685
Dh	783	AKKQKVEDVNDKKRAEVAAKKIDELKIEAEANI KTAEVA-----AKQKEX	832
Qy	666	DDE-----GGPTLGGDSHRAEESKQEKAGTAVPASIOEQUOQSSSPBPAPS	738
Dh	833	DEQLTLETVVSSKSAEKELEKQOIKKAAEADVAKQKELNKENKLEIAKASADKL	892
Qy	739	BEGEGVSTESPK--RLVTRPKKSKKLEEK-----AESVSEJOSTELBPSRBSWVI	790
Dh	893	KLEESAAKSKSVSESVKFGKEKTKAGKTVQVESSEPTSKTIDTVQVATEADETP	952
Qy	791	KKFIPGRKKRKAADQOQATVEDSGPV--EINEDDPNVAVPLSEYNV-EREKMEAG	847
Dh	953	KKKIIKKKTEKSDSISQKATSEKVSQKQEDDEPTKAV---SELQMTHEADSKKQK	1009
Qy	848	NTELPOLLGAVVSEBELSKTLVHTVSVAVIDGTRAVTSVEERPSMISASVTEPLEHTAG	907
Dh	1010	ETD-----EKLK-----LDAEIAKTKQEAD-----EKSKL	1035
Qy	908	EAMPPEVEVTEKOLIAETFEVLQTLPEGDADHDVMTSEVFTSEAYTATPSALTE	967
Dh	1036	DAOEKIKKVSDDAARE-----KELND-----KLLESEIATKASADKLE	1079
Qy	968	EVTESAGAEETDMVSAVSQULTDPSPTTEBAPTVEGESGVLDTE-----EER	1016
Dh	1080	EOQAKKAAE---VEAAKKQEKEDQK-----LDTAASAKKAAEKLLEK	1123
Qy	1017	OTOALIQAVD-----KVXESQVPATQTVQ-----RT	1044
Dh	1124	QAQIKKAGADAVAKQKELDEKKNLEANKKSAAGRKIEEESAASKQVTEBOAKLDAQT	1183
Qy	1045	GSKALEKEVEEDEVLAEBEKDVMFKPQV-----EKAHEHLAQSSTGATPES	1097
Dh	1184	KAKTAKQOTJLEKEKEKSTKSESSEFATVDEKPKKVLKKTTEKSDSISQKSTSTVSES	1243

Db	668	KKATEKSDSSISQ	SNVLKRA	DDKSKSD	VDVDS	SKTTEDDT	KVATVDS	KLEKA	DTI	--	725
Qy	567	EKGP	LEAF	PODEAB	EETG	TS	DG	KREB	GI	TP	MA
Db	726	KQI	ET	EV	VD	DD	KSKV	LV	--		
Qy	627	SAT	LS	ST	D	S	T	S	T	S	T
Db	783	AK	KO	EV	D	N	L	K	R	E	A
Qy	666	DDE	-----	GG	P	R	L	T	G	G	D
Db	833	D	E	O	L	E	T	E	V	S	K
Qy	739	E	G	E	G	V	T	S	E	S	F
Db	893	K	L	E	E	S	A	K	S	K	V
Qy	791	K	K	F	I	P	G	R	K	R	A
Db	953	K	K	I	I	K	K	T	E	S	D
Qy	848	N	T	E	L	P	O	L	I	G	A
Db	1010	E	T	D	-----	E	K	L	-----		
Qy	908	E	A	M	P	V	E	A	T	E	K
Db	1036	D	A	O	E	K	I	K	S	E	D
Qy	966	E	T	E	S	A	G	E	E	T	D
Db	1080	E	O	A	O	K	K	A	A	-----	
Qy	1017	Q	T	O	A	L	I	Q	A	N	A
Db	1124	Q	A	O	I	K	A	G	A	D	A
Qy	1045	G	S	K	A	L	E	K	V	E	E
Db	1184	K	A	K	A	E	K	O	T	L	E

QY 1098 LEVPEVTADVDHATCQVVKIQLQMEQAVPESSET--LTDS-----ETNGSTPLADSDT 1150
 Db 1244 -----AGPSESETQKVAADAAARKQKQETDEKQKLEAEIT 1275
 QY 1151 ADGTOODETIDSQDSKATAAVROSOVTEEAATAQKEEPSTLPNNVPAQEEHGEPEGRDV 1210
 Db 1276 AKSADEK-----SKLEASKLKAEEVAANKQKQEKDQLKLDT---EASAKAAAEK 1326
 QY 1211 LEPTQOELTAAAVPVL-----KTEVQGEVDWLQGEKVKKEQEVFVHSGPNSQK 1261
 Db 1327 LELEKQSHIKAAAEVDVAKKQKLEEKORLESEAAATKKAADAEKLKLEQ-----KKX 1378
 QY 1262 AADVTYDSEVMGVAGCOKESTEVSQSLSEGEOMETVKEKEKRETKPEQVSEGEQHTAA 1321
 Db 1379 AAETA-----LIEQKEQKLAQE--QSRLEDEAKKSAEKQKLESETSKQTEE-----A 1426
 QY 1322 PEHEGTYGKPVLTIDMPSSERKALGSLGGSPSLPDQDKAGCTEVSQVQSLDTTWTQ--- 1377
 Db 1427 PKE-----SVDEPKKKVLLKKTEKSDSSISQSKS 1457
 QY 1378 ---TAEAVE-----KVIVTWSISETGESPECVGAHLLPAEKSSATGGHWTLQHAEDTVP 1428
 Db 1458 AKSTVDAAEETLESDFNLVEKTKVQKQVQSPD-----ESTSATIKRDPQAQKTEEISK 1508
 QY 1429 LQPSQAESEPIIIVTPEPSTLHPDLQGEISASQRESEEDKPDG--GPDA----- 1478
 Db 1509 QDPGDEKTKTTTQDKPKPKPEDS-----EATPKKRVVKKTKQKSDSVASDASLADVSKLS 1561
 QY 1479 DGKESTAIEKVLK-----ASPEILELES--KSNKVLVNIQTAVDQFA 1519
 Db 1562 DDVEPKKKVLLKKTEKSDSVISSETSVDTIKPESVEIPTERAEQMLH-----NRFS 1615
 QY 1520 RTETAPET--HAYDSQTQ 1535
 Db 1616 -TDSAVESEPKNAHKDTE 1633

RESULT 7
 S37431
 N:Alternate names: ankyrin B, 440K splice form; ankyrin-B; brain ankyrin; non-erythroid
 N:Contains: ankyrin 2, short form
 C:Species: Homo sapiens (man)
 C>Date: 06-Jan-1995 #sequence, revision 06-Jan-1995 #text, change 13-Aug-1999
 C:Accession: S37431; A39643; B39643; A40334; A49462; S14533; S14569
 R:Chan, W.
 submitted to the EMBL Data Library, September 1993
 A:Reference number: S37431
 A:Accession: S37431
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-3924 <CHA>
 A:Cross-references: EMBL:226634; NID:9406287; PIDN:CAA81387.1; PID:9406288
 R:Otto, E.; Kunimoto, M.; McLaughlin, T.; Bennett, V.
 J. Cell Biol. 114, 241-253, 1991
 A:Title: Isolation and characterization of cDNAs encoding human brain ankyrins reveal a
 A:Reference number: A39643; MUID:91302466; PMID:1830053
 A:Accession: A39643
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-2077 <OTT>
 A:Cross-references: GB:X56957
 A:Accession: B39643
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1443, 3585-3924 <OTT>
 A:Cross-references: EMBL:X56958
 R:Tse, W.T.; Menninger, J.C.; Yang-feng, T.L.; Francke, U.; Sahr, K.E.; Lux, S.E.; Ward,
 Genomics 10, 858-866, 1991
 A:Title: Isolation and chromosomal localization of a novel nonerythroid ankyrin gene.
 A:Reference number: A40334; MUID:92009921; PMID:1833308
 A:Accession: A40334

A:Molecule type: DNA
 A:Residues: 463-474, 'PB', 477-495 <TSE>
 A:Cross-references: GB:M37123; NID:9178647; PIDN:AAA62828.1; PID:9178648
 R:Chan, W.; Kordeli, E.; Bennett, V.
 J. Cell Biol. 123, 1463-1473, 1993
 A:Title: 440-kD ankyrinB: structure of the major developmentally regulated domain and s
 A:Reference number: A49462; MUID:94075409; PMID:8253844
 A:Accession: A49462
 A:Status: preliminary; nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-3924 <RES>
 A:Cross-references: EMBL:226634; NID:9406287; PIDN:CAA81387.1; PID:9406288
 C:Genetics:
 A:Gene: GDB:ANK2
 A:Cross-references: GDB:127607; OMIM:106410
 A:Map position: 4q25-4q27
 C:Superfamily: ankyrin; ankyrin repeat homology
 C:Keywords: alternative splicing
 F:2-3924/Product: ankyrin 2, long form #status predicted <MAT>
 F:12-1443/3585-3924/Product: ankyrin 2, short form #status predicted <MA2>
 F:96-95/Domain: ankyrin repeat homology <AN01>
 F:96-128/Domain: ankyrin repeat homology <AN02>
 F:129-161/Domain: ankyrin repeat homology <AN03>
 F:162-190/Domain: ankyrin repeat homology <AN04>
 F:191-223/Domain: ankyrin repeat homology <AN05>
 F:232-264/Domain: ankyrin repeat homology <AN06>
 F:265-297/Domain: ankyrin repeat homology <AN07>
 F:298-330/Domain: ankyrin repeat homology <AN08>
 F:331-363/Domain: ankyrin repeat homology <AN09>
 F:364-396/Domain: ankyrin repeat homology <AN10>
 F:437-429/Domain: ankyrin repeat homology <AN11>
 F:430-462/Domain: ankyrin repeat homology <AN12>
 F:463-495/Domain: ankyrin repeat homology <AN13>
 F:496-528/Domain: ankyrin repeat homology <AN14>
 F:529-561/Domain: ankyrin repeat homology <AN15>
 F:562-594/Domain: ankyrin repeat homology <AN16>
 F:595-627/Domain: ankyrin repeat homology <AN17>
 F:628-660/Domain: ankyrin repeat homology <AN18>
 F:661-693/Domain: ankyrin repeat homology <AN19>
 F:694-726/Domain: ankyrin repeat homology <AN20>
 F:727-759/Domain: ankyrin repeat homology <AN21>
 F:760-792/Domain: ankyrin repeat homology <AN22>
 F:793-825/Domain: ankyrin repeat homology <AN23>

Query Match 6.1%; Score 493; DB 2; Length 3924;
 Best Local Similarity 20.4%; Pred. No. 5, 1e-10;
 Matches 422; Conservative 270; Mismatches 680; Indels 698; Gaps 99;

QY 68 GDVHVOENQEGOEVEVDVQGRESEVDREKDRVVEEMAANSTAVEDITKQGEETSE- 126
 Db 1497 GSIKVKELVKAABEPGPEFIVERVKEDL---EKVNEILRSGTCTRD-ESSVQSSRSR 1552
 QY 127 -IIEQ--IPASENNVVEVQPAESQANDVGFKKVFGFKFTVKDK-----NEKSDTV- 178
 Db 1553 GLVEEWEVTVSDEEIEEAAKQAPLEITEYPCVEVRIDKEIKGKVEKDSGLVNYLTDDL 1612
 QY 179 -----QLLVTKQDEGGAESVAGDHQPSVETAVGESASKES-----ELKQSTEK 225
 Db 1613 TCVPLPKEQLQVQKAGKKCEA-----LAVGRSSEKGGKIDIPDETOSTOK 1659
 QY 226 Q-----EGTLKQEQSSTEIPQAESDQAAEEAKDEGEKQEKETKSP----- 270
 Db 1660 QHKPSLGIKKPVRRKLKEKQKEEGLQASAKA--ELKKGSSSESLGDPGLAPEPLPT 1717
 QY 271 -SPSPVNSSET--TSSFKPKFTHGWAGWRKKTSTFKSKEDD-----LETAERKKEQEA 322
 Db 1718 VKATSPLEETPIGSIKD-----KVKALQKRVEDEQKGRSKLPTRVKGKEDVPKK 1767
 QY 323 VDEEEKETEPASEBQ-----EPASDTQCARLSADYKVE----- 357
 Db 1768 TTHRPHPAASPLKSERHAPGSPSKTEHSTLSSSAKTERHPVPSPKTEKHSFVSPS 1827
 QY 358 LPLEDQVGDLEASSEBKCAPLA-----TEVFDEKMEAHQEVV----- 394

```

Db 1828 AKTERHSPASSSSKTEKISPVSPSTKTERHSPVSSKTRHPPVSSGKTDKRPVSPSG 1887
Qy 395 -AEVH-----VSTVEKTEEO--GGGGAEGGVVVEGG--ESLPRKLAEP 436
Db 1888 RREKHPVPSRGREKRLPVSPSGRTDKQPVSTAGTEKHLPVSPSGTEKQRPVSPSTK 1947
Qy 437 QEVPOEAPAEELMKSRKNCVSGD-----HTQULTDLSPEEK-----TLPRHP-- 479
Db 1948 TERIEETMSVREIMKAFQ---SGODPSKHKTGILFEHKSAKOKQPOEKGKVRVEKEKGPIIL 2004
Qy 480 ---EGVSVEMLSQERIKVQSSPLK---LFSSGLKKLSGKKQKGRGGGDE--- 530
Db 2005 TOREAKTENOTIKRGORLPVTGTASKRGVRSIGVK-----EDAAQKKEKVL 2056
Qy 531 -----PGEYOHITESPES--ADEOKGSSASPEE----- 559
Db 2057 HKIPREVQVPEEHSREBEVPEKEMADQGDMDLQISFDRKSTDPSEVIKQLELDNDK 2116
Qy 560 -----PEETCLEKGPLAPQ-----DGEAE--- 580
Db 2117 YQOFLRSEET---EKAQLHDQVLTSPFNTTPELDYMKDEFLPALSLQSGALDGSSESILK 2173
Qy 581 -----EGT---TSOGEKKEG--TPMASFKKM--VTPRK---RVARPSGSDK 618
Db 2174 NEGVAQSGGSLMEGTPQISSSESYHGEGLAETPETSPELSFPKSKSEBQIGETKSTK 2233
Qy 619 EEBLEKVKY-----ATLSSTDSVSEMOEVTKEGEBQKPEEKREY 660
Db 2234 TETTTTIRBEKHPHTKIDITGSEERGATVTEDESETSTSFOKEATLIG--SPKOTSFKQD 2292
Qy 661 DTSVSWMALICVSSSKRKARKAS-----SSDDEGGPRTLGGDSHRAEASQKAGATD 713
Db 2293 D-----CTGSCCVALAKETPTGLTEBAACDE--GORTFGSSAKHTQDSEAEQ--- 2338
Qy 714 AVAPSTQEDQAGSSSPBA-----GSPSEHGVSFWSEFKRLVTPRKKS----- 759
Db 2339 ---STATSDETALPLPEASVKTGTGTESKPOGVIRSPQGLELAPSRSEVLSAVADD 2394
Qy 760 ---KSKLEEK--AEDSSVEQLSTEIEPS--REESWVSIKKIFGRKKRA----- 802
Db 2395 SLAVSHKDSLSEAPVLEDNSHKTPLSLBPSPLKESPCRDSLESSVEPRKMGKGTIPSH 2454
Qy 803 -----DKOEOATVEDSGVEINEDPN----- 825
Db 2455 PLPAVAKTELTTEVASVSRLLRDPDGSADSDLEQTSLMESSGKSPSPDPSSEEV 2514
Qy 826 -----VPVV-----PLSEYNVBERE--- 841
Db 2515 YEVTPTTDTVSTPRAPVHIECAEEDDSENGEKKRFTPEEEMFKMTKIKWDELBOEAKQ 2574
Qy 842 -----KMEAQNTLPTQLGAVVSEELSKTLVHT-----VSAVADIGTAVTS 885
Db 2575 KRDYKPEKQEESSSDPDADCSVDDE---KHTSGDESGVPVLTSESKEVSS 2629
Qy 886 VEERSSPSW-----ISASVTEPLEHTAGEMPVEEYTEKDIATERTVLI 929
Db 2630 SSSSEPELAKQKKGADSGLLPEPVIRVQPPSLPSSMDNSNP--EEVQFQPVVSKQ--- 2684
Qy 930 TOTLPBGKAHDMVTSEVDTSEAVTATETSEALRTE-----EVTENS 973
Db 2685 -YTFKMNEDTQOEBPGKSEEEKDESHLA--BDRHAVSTEAEDRSYDKLNRPDQPKCDGH 2742
Qy 974 GAEETIDMVA--VSQUTDPS-----DTTEATPVQVEESG 1007
Db 2743 GCAMPSSSARPVSSGLOPFGDDVDDEQPVITYKESIALQGTHEKOTEGSELVLSAESP 2802
Qy 1008 VLDTEEBERQOAILQ--AVADKYKEESQVPATQVQRTG--SKALEKVEEVED----- 1058
Db 2803 QADCPSESFSSSSLPFCIVSEGEKLEDEDISATSSIOKTEVTKTDFEFENLPDCCSODS 2862
Qy 1059 -----SEVLASEKKEKVMKGPQOEGAEHLAAGSEFGQATP-----E 1096

```

```

Db 2863 SITTODRFMSDVPVSLDAENDEIYDQITSFVENVPSQSFSSSEESKQTDANHTTSFH 2922
Qy 1097 SLEVPBV--TADVDDHATC-----QYIKLQQLMEQAVAPSSSETLIDSE 1138
Db 2923 SSEVSVTITSPVEDVVAVSSSSGTVLSKESNFEQODIKMESOLESTL---MEMQSDSV 2978
Qy 1139 TNGSTP--LADSDTADQTOQDSTIDSQ---DSKATAAV-----OSQYTEE---AATA 1184
Db 2979 SSSFEPTMSATTYVGEQISKVITTKTDVSDSWSSEIHEDAEABAKVEEBOQIFGLMV 3038
Qy 1185 QKEEPTLPNNVPAQOEHEEPEGRDVLPTQOEL--TAAAVPLAKT---EVQGEVDM 1239
Db 3039 DRQSGTTPDPTTPAR-----TPTEGGPTSPQNFPLQOEGELFMTSSGAID- 3085
Qy 1240 LDGEVKEEQEVFVHSGNSQKADVTYDSVM--GVACQCKE--STEVQSLSEBEMET 1297
Db 3086 MTKRSYADESFHFQIG---QESREETLSEBVKEGATADPLPLETSAESIALSESKEV 3142
Qy 1298 DVEKEREKTEKPEQVSEEBEQ-----ETAAPEHGTGKPLVLTIDMPSE 1341
Db 3143 D---DEADLLPDSVEEVEELIPASDAQNSQMGISASTETPKAV---SVGTQDLPTVQ 3196
Qy 1342 RGKALGSLG--SPSLPDQDAGCIEVQVQSLDTTQTOTAABVKEVLETVVISERG--ESP 1398
Db 3197 TGD-IPPLSGVKQISCPDSEBA---VQVQDLFSLTRS-----YSDRGDDSP 3241
Qy 1399 ECVGAILLPAEKSSATGGMWTLQHAEDTVPLCP-----ESQAEISPI---ITPAPE 1447
Db 3242 DS---SPEEQKSVI-----EIPAPMENNVPTESKSK-IPVRTPMTSTPAP 3284
Qy 1448 STLHPDLQGEISASGRSESEEDKPDADGKSTAIKVL--KAPBEILLESKSKI 1506
Db 3285 SA---EYESVSEDFLSSVDEBNKAD---EAKPKSKLPVKPQLOVEQQLSDLTISVOKT 3338
Qy 1507 VL---NVLOTADQFARTETAPETHAYDSQTOVPACRLDREPNRCMKAKMAKHPV 1562
Db 3339 VAPQGDMASTIAPDRKSES--DASSLDSKTKCV-----KRSVTEETESREK-- 3387
Qy 1563 PQRPREDQVLTVLEMAQPRKCLPRLQKA 1592
Db 3388 ---AELE-LESEEGATRPKILTSRLPVKS 3413

RESULT 8
T29757
protein UNC-89 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 03-Dec-1999
C:Accession: T29757
R:Du, Z.; Le, T.T.; Wilson, R.
submitted to the EMBL Data Library, May 1997
A:Description: The sequence of C. elegans cosmid C09D1.
A:Reference number: Z20679
A:Accession: T29757
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-6642 <DND>
A:Cross-references: EMBL:AF003131; PIDN:AB54132.1; GSPDB:GN00019; CESP:unc-89
A:Experimental source: strain Bristol N2; clone C09D1
C:Genetics:
A:Gene: CESP:unc-89
A:Map position: 1
A:Intons: 17/2; 108/3; 154/2; 211/2; 265/3; 326/2; 352/3; 426/2; 454/1; 500/1; 537/1; 6
3/3; 5917/1; 6027/1; 6061/3; 6153/2; 6515/1; 6552/3; 6609/1

Query Match 6.1%; Score 489; DB 2; Length 6642;
Best Local Similarity 21.8%; Pred. No. 1.3e-09;
Matches 348; Conservative 234; Mismatches 670; Indels 346; Gaps 68;

Db 1132 TYQKRPAPBEPFLRVSLVTEKGEAVFSAHAFGLP-----LPTYEM---SYNG--- 1177

```

```
QY 67 QGDVHVQENQEGEEVVDVCGRESEDEVREKD-----RVEEMAANSTAVED 115
Db 1178 -----RKVRDQGEARGVTRDSTVDGASILTIDTATYSEVNNHLTISVVAENTLGAEE 1230
QY 116 ----ITKDGQEBTSEITIEQIPASNNVVEEMVQPAESQANDVGFKVFKFVGFKF--TVKK 169
Db 1231 TGAQLTIEPKKE-SVVVEKQDLSSEVQKIEAQOVKEASPEATTTIMETSLSTKTTM 1289
QY 170 DKNEKSDTVOLLTVKKDGGGABASV----GAGDHQPSVETAVGSAKESBLKOSTEK 225
Db 1290 STTEVTSTVGVTETKESSEATTVIGGSGGVTEGSIIVSVKIEVVSKTD---SQTDV 1346
QY 226 QEGTLKQEQSSTEIPLQASDAQAEBAKDEGEKQEKPTKSPESPSPVNSSETSSPK 285
Db 1347 REGTPKRRVSFAEELPKVIDSDRKKKSPDKKSPDKKSEKTEEPASPT----- 1397
QY 286 KPFTHWAGWRKKT--SFKKSKK--DDLTAETAKKEQAEKV---DEBEKEKTPASEEQE 339
Db 1398 -----KKTGEVKSPEKSPASPTKKEKSPAEEVKSPTKKEKSPSPPTKKEKS 1446
QY 340 PAEDTQARLSADYKVELPLEDOVGDLASSEBKCAPLATEVDFDEMAHQVVAEVHV 399
Db 1447 PSSPT---KKTGDEVKSPKPT--KKEKSPKPKPDVKSPVKKEKSPDATNIVEVSE 1501
QY 400 STVEKTE----EEOGGGAEAGGVVVEGTGESLPPKLAEPQEVQEAPELAKSRE 434
Db 1502 TTLEKTETMTMETHESESRSTSVKKEKT-----PEKVDKPKSPTKKDKSPEKSITE 1556
QY 455 MCVSGGDHTQLTDLSPREKTLKPHPEGIVSEVEMLSQERI-KVQGSPLKKLFSSSGLKK 513
Db 1557 I-----KSPVKK--EKSPEKVEKSPASPTKKEKSPKSPASPTKK---SENEVK 1599
QY 514 LSKKKQKRGKGGGDE--EPGEYQHIHTESPESAD-----EQKGESSAS- 555
Db 1600 SPTKKEKSPKSVVEELKSPKE-----KSPKADDPKSPPTKKEKSPKSPKSPKSPK 1653
QY 556 ----SPEEPE----TTCLEKPLEAPQDGEABEGTSDGCKREKREGITPWASPKQWTPK 607
Db 1654 KKEKSPKVEKPTSPPTKKESSPTKKTDD--EVKSPPTKKEKSPQTVVEKSPASPTKKEKSP 1712
QY 608 K----RVRRSESDKBELEKKSAT---LSSTDSTVSEMQDEVKTVGEBQKPEEPKRRV 660
Db 1713 KSVVEEVKSPKSPKAEKPKSPPTKKEKSPKSAAEVKSPTKKEKSPKKEKSPKAEKPKS 1772
QY 661 DTSVSWEALICVGSKKRARKASSDDGGPRTLGGDSHRAESKDKAGTDAVPASTQ 720
Db 1773 PT-----KKESSPVKMADE-----VKSPTKKEKSPK---VEEKPASPT 1809
QY 721 EQDAQGSSSPAGSPSEGEVSTWESFKRLVTPRKKSXKLEKAEKADSSVQOLSTEIE 780
Db 1810 KKEKTPKSAAEELKSPKKE-----KSPSPPTKKTGDSKESKPE-----K 1851
QY 781 PSREESNVSIKFTIPGRKKRADGKQOATVEDSGPVEINEDDPNPAV--VPLSEYNVAV 838
Db 1852 PEKPKSPKPKKPPGSPK-----KKSPEAEKPPAPKLTDLQTVNKTDLAHFEV 1907
QY 839 EREKMEAQGNTELPQLIGA--VYVSEB-----LSKTLVHTVSVNAVIDGTRAVTSVEE 888
Db 1908 VEHAETCKWFLDCKEITTAOGVTVKDDQDFEFRCSIDTTFGSGTVSVV-ASNAAGSVET 1966
QY 889 RSPSWISASVTEPLEHTAGAMPVE-----EVTEKDIIAETPVLTT-----QTLPEKQDA 939
Db 1967 KTELKVLTEPKETKPKBFTDKLRDMETVTKGDTVMQDVIALLHSPLYKWYQGNLLEDGKNG 2026
QY 940 HDDMVTSEVDFTSEAVTATETSEALRTEEVTEASGAETITDMVSQVSLTSDPTTEAT 999
Db 2027 --VTIKNEENKSLIIPNAQDSGKITVEASNEVGSSSS-----AQLTVNPPST---T 2074
QY 1000 PVOEVESGLVDTBEERQOTQAILQAVADKVEKSOVPATOTVORTGSKALEKVEEVEEDS 1059
Db 2075 PI--VVDGPKSVTIKETETAEFKATI-----SGFPA-PTVKWTIN-----EKIVVES 2118
QY 1060 EVLASEKEKDV----MPKGPVQEGAEHLAQGSETGOATPESLEVPVETADVHVATCOV 1115
```

RESULT 9

T15348

hypothetical protein B0350.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C:Accession: T15348

R:Gattung, S.

submitted to the EMBL Data Library, February 1996

A:Description: The sequence of C. elegans cosmid B0350.

A:Reference number: Z18332

A:Accession: T15348

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-5170 <GAT>

A:Cross-references: EMBL:U50071; NID:gi208871; PID:gi208877; PIDN:AAA93447.1; CESP:B0350.1

C:Genetic: CESP:B0350.1

A:Introns: 48/1; 5039/3; 5116/3

Query Match

6.0%; Score 488; DB 2; Length 5170;

Best Local Similarity 20.0%; Pred. No. 1.1e-09;

Matches 391; Conservative 249; Mismatches 678; Indels 634; Gaps 89;

QY 5 SSTEQRSPQAGSDTTPSELVLSGHGPAEASGAAGD----PADADPATKLPKNGQLSS 60

Db 818 AESERQVPSPVSHQYEPHVVETTTTTNTVTSNIYDDEDNVPSSDEPATQHQQ----S 872

QY 61 VNGVABQGDVHVQENQEGEEVVD-----EDVQRESEDEVREKD 101

Db 873 ETSVHRSHDPVSVEESDGEGLGKVLGFAGKAGVAGVVAAPVALAAVAKAAYDAFEK 932

QY 102 RVEEMAANSTAVEDITKQGEETSEIIQIPASENNVEEMVQPAESQANDVGPK----- 155

Db 933 -----DEDDTSHSPES-PVPEYQ-SEQYQDSDSAQSSHTDFEHMPES 973

QY 156 -----KVFPVGVGPKFTVKDKNEK-----SDTVQLLTVKK---DEGEAEASVGAGDH 200

Db 974 PIEHEKETBEFDHSHPSVPVLSKEKREHQVTSITTTTTTITVREYNDEPDEQ-----EH 1027

QY 201 QEPS-----VETAVGESAKESSELKOSTEKQEGTLKQEOSSTETPIQAES 245


```

Db 1028 OCPHSPASSTHTEAHPIVTTTTTTRERQ----EEPELEKYKQEDNSKRKSPSSHSQ 1083
QY 246 DQAAEE-----FAKDEGEKQKEPTKSPESPSPVNSE-----279
Db 1084 ENLVETTTTTTTVTSEYDEPEHEFEQETGKSPASSHVSESRQVLSPVSDPRHVAE 1143
QY 280 --TTSSFKKFFPHGMAWKKTSFKKSKE--DLETAKEKKKEQEAKEVDEEKKTEPEASE 336
Db 1144 TTTTTLVTRQF--HDESEKSDSPNRPVVAQSTHSSNIESHQOFSEKDDSQKSPFVKS 1202
QY 337 EOPPAE-----DTDOARLSADYEKELPLE-----361
Db 1203 EDEPVKHSYKRETSTETTEREPRDLSEBRS--EPAQSPHEVSPITTEENIVAPSSS 1259
QY 362 ----DQVDDLASSSEKCAPLAT--EVFEKKEAHQEVAAHVAVSTVEKTEEBQGGGE- 414
Db 1260 VKSEYSEGHVPSVJETKTTTTRREFYDDQDEHENQTOSEELRASSIPTEEHEHGHSL 1319
QY 415 -----AGGVVVGTSGLPPEKLAERQVPO--441
Db 1320 FKETTTTTTTRFYPDEPNVEELQOPSPAPSSHVSEGHASESP--VAQOQETPQTR 1377
QY 442 ---EAPPAELMKSRMCSGSGDHTQTLDSP---EKTLPKAPGIVGEVEMLSQERL 495
Db 1378 EFHEDSPAAQYFHEEY-----ESHVLTEQAPLITQOQHP--SQDESDGEGLS----1426
QY 496 KVGQSPKLKLFSSSG-----LKKLSGKKQKQKGGGDEBEGE-----YOHHTESP 542
Db 1427 KVLGFAKKAQWAGVVAAPVALLAAGAKAAYDALKKDDDEEDQERESLIRGERSIDSP 1486
QY 543 ESAD---EKGESSASSPBEPE---ETCL-----EKGLEAPQDGEABEGTT 584
Db 1487 HASEQOIEEHEHFEBSVPSEKHVETTTTTTTRREFPDEHPL--VSQLEGGKSGP 1545
QY 585 SSGEK-----KREGITPMASFYKM--VTPKKRRPSESDKEEEL 623
Db 1546 AUSEKSLPHVETTTTTTTRREFDKNDSPVSEKEQERTYSRVEYTAEDDEBEH--1603
QY 624 KVKSATLSIDTSTVSEMODEVKTVGEQKP--EPPKRRVDTSVSMALICVSSKKRARA 682
Db 1604 ---HYETTTTTTTRKEVIDDSQEMGDDDBKQSPQVET-----TTMTSREX 1650
QY 683 SSDDGCGRTLCGDSHRAEASKDEAGTDV-----PASTQODQAQSSSPBPAG 735
Db 1651 DNDDE--TRSEAGDSHITETKT-----TTVREFHSGQPEETETDEVE-1695
QY 736 SPSEGGVSTWSPFKLMPRKSKSLBEKADSSVEOLSTIEERSREBSWVSIKKFIP 795
Db 1696 -----LPPKLEEDNVSEYSSST--SVSRVFRD--EPHITETTTT 1735
QY 796 GRRKKRADQKQATVEDSGPVEINEDPNVAVPLSEVNAVEREKMEAQNTLPLL 855
Db 1736 VIREHNEDEETDDQKAAPIFSQ-----EHQDDDSQASHQDHRSP--1780
QY 856 GAVVYSEELSKTLVHTVSAVIDGTAVTSVERKSPSMISASTPELLEHTAGAMPVEE 915
Db 1781 -----VESEKSVKHT-----TETTT-----TTVTRQLYD-----DE 1807
QY 916 VTEKDIABETPLTQTLPEGKAHDMVTSYDFTSEAVTATESEALRTEVEITASGA 975
Db 1808 ABE--IRGESPVATE-----EHEHVSSTKSD--ESQHVPSV 1841
QY 976 EETTDWVASVQTLSDPP--TTEATPVQEVESGVLDEEERQTOAILQAVADVKAEE 1033
Db 1842 IETTTTTTTRREFYDDQDELQREDHTQSEBRRSI--PTRETHEEDHLIK-----1890
QY 1034 QVPATQTVORTGSKALEKVEVESEVL--ASEKEKV--MPKGPVQZGAHILAQSGSET 1090
Db 1891 ETTTTLVTRREFDEPENKLDQSFSLSPSSHVSESIIVPESPV--AKQCHIPQTRER 1948
QY 1091 GQATPESLEVPETAVDVHATQOVITKQOLMEOAAPRESSLTLTSPNNGSIFPLADSDT 1150
Db 1949 HEDSPAAYFHE--DEYEHQVPTQAPL--LTEQOHQSPESGEE--SPGEGFGSVLGPAAKK 2003

```

```

QY 1151 A-----DGTQDETIDSQDSKATAAVRQSYTER-----1179
Db 2004 AGWAGGVAAAPVALLAAGAKAAYDALKKDEDEQEE-----RESLIREBSFDSPHA 2057
QY 1180 -EATAQKEEPESTLPNNVPAQEBEGEE---PGRVLPEPTQOELTAAVAVLAKTEVQGEQ 1235
Db 2058 SEQOIEKENKRFESPPVSEKHHQSSALPOESVSQIIEK-----SRFNDSE 2107
QY 1236 EYDWLDGKVKVEEQEFVHSGPNSQKADVTYDSEWGVACQEKESTEVOSLSLEGEM 1295
Db 2108 EFGVSDHYTEDDQESLSPKESGPAFSQFTSEKE-----QDRSDPISH-----Q 2153
QY 1256 ETDVKEKEKRETKPEQVSEGEQETAP--EHEGTGKRVLLDMSSSRGALG-----S 1348
Db 2154 KEDISQFQNESSPEVQKSEOPDEKPLERQGSYSSG---SPKSPGSGITTGDEEKA 2209
QY 1349 LGG-----SPSLPDQKAGCIEVQOSLDT 1373
Db 2210 LSGVPEPEDRPENFASHEKTEATSDENLPESDKAPASP--VPSSDSN---RVLETFT 2264
QY 1374 TVTQTAEAVER-----VETVVISETGESPECVGAHLIPAKS-----SATGHWTLQ 1421
Db 2265 TTTVTRHEPEPEDHSDYVESQEVSSSG--SP-----VPSEKSVDRVLETFTTTTTR 2316
QY 1422 H--ADTVPLGPE--QAESIPITVPAP--ESTLHPDLQCEISAS-----1461
Db 2317 HFEADEIPITVSESHDDQAA--PVPSEEDVHQQIOTTTTTRVTKHFVFPDEI 2370
QY 1462 QRESEEDKPDAGPDADG-----ESTAIKYLKA--PE-----LDEL 1499
Db 2371 DSEHNNESDKIAGSPVPSEEDSGRVETTTTTTTRHFEHEDHSPVVOQEYSASES 2430
QY 1500 ESKSNKIYLVNIQTAVDQFARTETAPETHAYD-----SQTVPACRLDSREPNRC 1549
Db 2431 PVPSEKSVARVLETFT-----TTVTRHFEDEHIILOGQESDSQISESITSNMR- 2484
QY 1550 WTKMDAKKHPVPOPRFDLOVLYLEMAQP 1581
Db 2485 -----ETSSSPVQSNRDEEFLPALAIAPYKOP 2510

RESULT 10
T26216
hypothetical protein W06A7.3c - Caenorhabditis elegans
C1:Species: Caenorhabditis elegans
C1:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C1:Accession: T26216
R1:Ainscough, R.
submitted to the EMBL Data Library, August 1996
A1:Reference number: Z20173
A1:Accession: T26216
A1:Status: preliminary; translated from GB/EMBL/DDBJ
A1:Molecule type: DNA
A1:Residues: 1-2484 <MIL>
A1:Cross-references: EMBL:Z78066, PTDN:CA851467.1, GSPDB:GNC0023, CESP:W06A7.3c
A1:Experimental source: clone W06A7
C1:Genetics:
A1:Gene: CESP:W06A7.3c
A1:Map position: 5
A1:introns: 1827/1, 1866/3, 1963/3, 1990/3, 2262/1, 2288/1, 2339/2, 2463/2

Query Match 6.0%, Score 485, DB 2, Length 2484,
Beet Local Similarity 20.2%, Pred. No. 5, 8e-10,
Matches 362, Conservative 244, Mismatches 648, Indels 538, Gaps 69,

```

```

QY 245 SD-----QAEEBAQDEGEKQKEPTKSPSPSPVNSETTSPFKKFTTHMAQWR 296
Db 59 GDCERGNIKISANENSVEPDGADKPAH--TEQELPTIILPESETTVT-----102

```


Qy	297	KKTSFFKSKBDDL-ETAERKKQOEAKVDE-----EKEKTBSPASBEQPAEDTD-QAR	348
Db	103	-----QLEDHLSETERKQNEVESIKSEKNVIGLEDSGDTRDTSGLAPAEADNEAT	155
Qy	349	LSADYEKVELPLEDQVGLSEASEKCAPLATEV--PDEKMEAH-----	390
Db	156	ITTDf-----VPLE-SAGDIPSGNEIKEVASPDVVGLEYIIGNINAPVVDNDIPNVFT	210
Qy	391	QEVVAEVHSTVBKTBEEGGGGEAGGVV--EGTGESL-----PPBKLAEPQOE-VPQEA	443
Db	211	PEVANDETVETFSVTAE-----EASIPVVVELEPIGDEYEFQRPVENFSEPSDNINLEE	264
Qy	444	EPABELMKREMCVSGDHTQTDLSPEBKTLPKHPGIGVSEVEM-----LSS	491
Db	265	SGAEQVLLLENM-----FTPLDELEPOQKILINEAKAHI--EIEASGDGETIKOHPPEVEN	316
Qy	492	QERIKVQGSPLKLFSSGKLGKQKQKRGKGDEEPGEYQHITHTEPE-----SA	545
Db	317	EQNVQESRVEIVHSFIGLES-SG---VGLTGAVSDSVANNVKE-NTESPDIISLEASG	371
Qy	546	D-----EOKGBESSASSPEEPETTCI-----EKGP	570
Db	372	DELSKLVEAREIITISKDAYSTDVPSRKTVSDVIGLEAGDKIVSNVSNVMGNPDQSQ	431
Qy	571	LEAPQGEABEGTTSQCKKREGITPWASFKNWTPKRVRRPSESQKE-----BELEKVK	626
Db	432	LALQEQADNVPELPIENSEQETVAVKATESCDHVVDQSQKTLBRASLEEDIMSPEVLGITS	491
Qy	627	SATLSSTSDTVSMQDQVKTGVEEQPEPKRRVRDTSVSWEALICVGGSSKKRARKAKSSD	686
Db	492	SQTLSDYLPVISEDQDSIPV-----PVTVEZE-----TSEKLVKASLEE	531
Qy	687	DEGGPRTLGGDSHRABEASKQEAAGTDAVPASTQBQDOAGSSSPGAPGSPGEGVSTW	746
Db	532	DVVSPEVLELDD---RVQNKNPSESVTADVASKTEGD--FSDSPDGRAT-----	575
Qy	747	ESP-KRLVTPR-----KKSXSKLEEKAEEDSSVEQ-----LSTEIPEPSRE-----	784
Db	576	ETFMKLVTVTENLPLAGDKLSEERIQEIRENETISQPKGSEDDLENANDPDDETIVEKI	635
Qy	785	-----ESWVSIXKFI-----PQRKRKADRGQEQATVEDSGPVEINEDDPNPVAVVPLS	833
Db	636	VSMAESLPIEAVISTEDGGTSDQPAQNAIPDSEETTVDDSQTEHFTDDN-----VKKS	690
Qy	834	EYNAVERKMEAGCNTELQPLGAVTVSBELSKTLVHTVSVAVIDGTRAVTSSVEERSPSW	893
Db	691	KENT-----PKAENDTEINYLPGGEGEPD-----NAEKRNAEAVSPND	728
Qy	894	ISASVTEPLEHTAGEAMPP-----VEEVTBKDIASETPLVTQTLPEGKDAHDDM	943
Db	729	ETSEIKQDLNLENGASGPDNNVQVDEAAQEDPTQPTVDETTSKISENMPKAPDTEDDN	788
Qy	944	VTSEVDPFTSAVTAETSEAL-----RTEEV	969
Db	789	ATBEIE-----AGLETFEAFGDAEHVSYLDANIEKLVAMADEPLPVDELVSIEERPEEV	842
Qy	970	TEASGABETDMDVSAVSQLTDSPTTTEETATPVQEVESGVLD-----TEEBERQTOAILQA	1024
Db	843	APAESTGEDEDI FRDRDRVTSLTGTQGNAPIQVIFVGGDGDNDPANADQERTSEHNELI	902
Qy	1025	VADVKVESQVPATQTVORTGSKALEKVEEVEDSE-----VL	1062
Db	903	ESDKSEEAITKNEEDVDQPTQSEBPLTSQGBESSIGNKIVAVVGVSVLGGAVIPGVCL	962
Qy	1063	ASEKEKQVMPKGPVQEGAGAEHLAQGSSETQCATPESLEVPEVTADVDHV--ATCOVILQOQ	1120
Db	963	ASNENEDAHADREVEET-----GDSTRDRPEEETFFSKLTSVMVENILPSTNDENPEAV	1015
Qy	1121	LMEQAVAPESSEITLTS-ETNGSTPLAD-----SDTAGTQQDETI-----	1160
Db	1016	SMVENVLVPVNTGDLGDSKEDNDPAPTAEAHSGEKNLRNDKTTDTRERDPBETITINKLVE	1075

Qy	1161	- - - - -DSQS KATAAVR QSQVT EEEAATQAQKE-	- - - - -PS	1190
Db	1076	NALPTGVGTSETSAPDAQELTETVVDHAGNDTSEYDAPEKSAAGGTVIEKFTSMIES	:	: :
Qy	1191	TLPNNVPAAQEHBGE-- --PGRDLVLEPTOQLTAATAAAPVLAKTEVG---	OQE	EVDWL
Db	1136	ILPVQAFTPENARSHIDQTGGASEIKDDQNPPESAEHCCKEVEASAPPQESAEPQL	1195	
Qy	1241	DCEKVKEBQVFVHGSPNSQKAAD-----VTYDSEVMGVACQEKESTESTVSLSLEG	1293	
Db	1196	EAKKDQDKETI- ----ENSEDAKKETVMKLVSLENILPVEAVLPSDST-VTKNSEDKK	1249	
Qy	1294	EMETTVEKEKETK----PEQVSESREGOETRAPE-----HECTYGKPVLTLDMPSERCK	1344	
Db	1250	ELETQ-EUSSKEIKTGSGQFYYPETSEA FSDPDEIFQRVKRASSTEP---KTQKTEPHA	1304	
Qy	1345	ALGLSGLGPSPLPDQBKAGCIEVOGSLDTTVTQTAEAEEKVIETVVISSETGSPECVGAH	1404	
Db	1305	PFIIVGSTEDDEQSIANVIDELVHEDD-----EKKYPEVTANTISVASENIDISTTAN	1358	
Qy	1405	LIPARKSATGGHWTLQHAEDTVPLGPSSQAESIFIPIVTAPAESTLHPDLOGEISASOR-	1463	
Db	1359	AVP--KTEVSSELQVATVEFELESAPBEESAIIPEVGEPLEKVEQPDLSONSPAPHKI	1416	
Qy	1464	-- -- -- -- --ERSEEEDRPDAGDPADGKESTAIE:	: : : :	
Db	1417	IDLHFNPDKDHEDYGN DYVPFGTESSESQADGNQENOEEDVVAELNFHFIROWRDED	1476	
Qy	1498	-- -- -- -- --KVLKAEPRILELESKNKIVLMNV	1510	
Db	1477	VISLOSLSLVAEVCITDVDASDVNEQDBESTLKILKVVPESEPILLSLEDFTNDPKVIHV	1536	
Qy	1511	-- -- -- -- --IGTAVDQFARTETAPTHAYDSOTQVPACKLDLRPNRCWT	1551	
Db	1537	PIPLMEPATMYLEEMVEMIADVAVESEM EVTTSEISEMAPQVSESTCP IEP----	1591	
Qy	1552	KMKDAKM-----KHFPVQPREDLVOL-TVLEAWAQPRKC LRLQLKAPVSK	1596	
Db	1592	-LAIDLKLFVEDDEKTPPEVPVPGVQVRRIPIEV EQAPTIPQRP RPAPKSE	1642	
 RESULT 11 T26215 hypothetical protein W06A7.3a - Caenorhabditis elegans C:Species:Caenorhabditis elegans C>Date:15-Oct-1999 #sequence_revision15-Oct-1999 #text_change15-Oct-1999 C:Accession:T26215 R:Ainscough, R. submitted to the EMBL Data Library, August 1996 A:Reference number:D20173 A:Accession:T26215 A>Status:preliminary; translated from GB/EMBL/DBJ A:Molecule type:DNA A:Residues:1-2607 <WIL> A:Cross-references:EMBL:Z78066; PIDN: CAB01522.2; GSPDB: GN00023; CESP: W06A7.3a A:Experimental source:clone W06A7 C:Genetics: A:Gene:CESP:W06A7.3a A:Map position:5 A:Introns:1827/1; 1866/3; 1963/3; 1990/3; 2262/1; 2289/1; 2412/1; 2462/2; 2586/2;				


```
QY 140 EMVQPAESQANDVGFKKVFPVGFK-----FTVKKDKNKSSTVQLLTVKVD 186
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 396 DKIERDK-----BFEFGKDEYFTQMEGGAGYITVEHGGQKRLTLT----- 436
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 187 EGGAEASVAGDHQEPS-----VETAVGESAKESL-----KQS 222
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 437 --EGVGEYKVIEMEQTKLDFLLPTNEKSDDFVSEMEERLFFNVSVMEBVKDKMF 494
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 223 TEKQEGTLKQ-----EQSSTEIPLOAESDQAAEEAKADGEBKQKPTKSP 269
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 495 IREQGTETEEVIHEDLSDEKIGIEKKELETFFQDVSDVQLDEQLIEEKLTEEEKLTEE 554
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 270 ESDSSPNSETTSFKFFTHGAGWPKKTSFKKSKEDDLLETAKRKQBAEKVDEEEK- 328
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 555 EKLTEE-----EKLTEEEKLTDE-QLTIEEKLTEBEKLTBEKLTBEKL 593
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 329 EKTEPASEOEPAED--TQOARLSADYKVELPLEDOVGDLSESSKCAPLATEVFE 385
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 594 TEEKLTEEKLTEEEKLTBEKLTBEIEFIE--EDKI-----SPKVEEPL-EDVIME 644
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 386 KMEAHQEWAEVHSTVEKTEEEGGGGEAEAGGVVVBGTGESLPPEKLAE----- 435
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 645 KMEKEYKEILLHEGKSVVTDQWDEKLLSENLEDDLIPELIEELVDEELVEEVAEEIVEE 704
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 436 -----POEVPOEAPABELMKSRMCVSGGDHTQLTDLSPBEKTLPHKEGIVS 484
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 705 FVVEVEITEEIKPELVEEVMPEED-----VDEVVPEE--LVPEPDSWE 746
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 485 EVELMSQOEIRIKVQSPKLKLFSSGLKLLSGKKQKRGGGGDEEPGE-YQHHTESPE 543
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 747 EVTTEELTEEV-IPPELIEKVBEELVEKVVAQKIV-----EVIPEELIEEVIETKP- 797
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 544 SAEQKEGE-SASASSPER-----PEETTCLEKGLPEAPOGEAEGETTSDGEKKREGITP 596
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 798 -ASEVPELVEEVIPEEIVPEEVPV-KEVPEEVEEVIPEEVEEVEEVEEVEEVP 855
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 597 WASFKQWTPKRVRRPSESDEKEELBKVKSATLSSTDSTVSEMQRDKVTVGBEQKPEEP 656
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 856 ----EELV-----EEMKPEEVPFSEVISEEL-----VEEMKPE--EWEVEVTE-- 892
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 657 KRRVDTSVSWEALICVSSKKRKAKSSDDGEGPRTLGGDSHRAEASKDKAGTDVAP 716
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 893 --KVVEEVPPEELV-----EEMKPEEVEEVEEVP 918
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 717 ASTQEO---DQAGSSSPPEAGSPSEGEVSTWBSFKRLVTPRKK-----SKSKLEEKAE 768
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 919 EELVEEVIPEEVEEVIPEEVEVHEAKHE-----ELFEKLV-PEVEVEEIIPEKELVEIPE 972
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 769 DSSVEQLSTEIPSR--EESWVS---IKKFIPIGRKKRA---DGKQE---QATVEDSGPV 817
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 973 EVLPPEELIEEVIPEEIEEVEEVIPEEVEEVIPEEIVQEVIPEDLMEEVPVEEIEEKIPE 1032
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 818 EINEDDPN---VPAVPLSEYNAVERKMEAGNTELPOLLGAVYSEELSKTLV----- 869
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1033 EIVEEEIHEEIVEEIIIP-----EEVVEEVPFAEVEEMIPKEIVEEVIPEELVEEVPVEVL 1088
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 870 -HTVSVAVIDGTRAVTSVEERSPSWISASVTEPLEHTAGAMP--VEEVTEDIIAEET 926
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1089 EEVVPEVLEEVIPEEVEEVIPEEVEEVIPEEVEEVIPEEVEEVIPEEVEEVIPEEVEV 1146
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 927 P-VLTQTLPGKQADDMVTSEVDFTSEAVTATSETSEALRTEEVEASGAETTTDMVSAY 985
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1147 PEVVEEVIPE--ELVEEVPVEL--LEEVISEELLEVEEVIPEELVEEVPVEEVEEMKPE 1201
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 986 SOLTQD-SPDT-TEBATP---QOEVESGVLDTTE-----BERQ 1017
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1202 EVLBEVPEEVLVEEVIPEELVEEVIPEEVEEVIPEEVEEVIPEEVEEVIPEEVEEVEQ 1261
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1018 TOAILQAVADKVEE-----SOVPATQTVQRTGSKALEKV--EEVEED-----S 1059
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1262 EQLVEEPIPEEVEEVLPEEVEEVIPEEVEEVIPEEVEEVIPEEVEEVIPEEVEEVEE 1321
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
QY 1060 EVLASEKEKDVMPKGPVQAGAEHLAGQSETGOATPESL---EVPEVTADVDHATCQVI 1116
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1322 EVIPEELVEEVIPEELVAEVIPEELVE---EVIPEELVEEVPVEEVEEVEEVEE 1376
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1117 KLOQLMEQAV-----APES--SETLTDSETNGSTP-----LADSDTADGTQODETI 1160
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1377 IPEELVEEVIPEVLVEEVEEVEEVEEVIPEELVEEMLPEELIEVVPVEEVEEVEEVL 1436
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1161 DSQDSKATAAVROSQVTEE-----EAATAQKEEPTLNNNPAQEEHGE--PGRDVL 1211
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1437 EE-----VPEEVEEVIPEELIEELVPEVEIEEILP-----EELIEELIPEEIE 1482
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1212 EPTQOELTAAAVVLAKTEVQGEVDWLDGEKVKEE-----QEVFVHSGNSQKAAD 1264
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1483 EVIPEELAEVIGEIVEEVEEVEEVEEVEEVEEVEEVIPEEVEEVEEVEEVEEVEE 1536
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1265 VTVDSEVMGVAGQCEKSTEVQSLSEEGEMETDVKKEKRETKPEQVSRBGEQETRAPEH 1324
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1537 EVIPKEI-----AEVLSAQVFEEVIPKEIVEEVEEVEEVEEVEEVEEVEEVEE 1586
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1325 EGTYKRPVLTLDMPSSERKALGSLGSPSLPDQDKAGCIEVQVQSLDITVTQTAAVEK 1384
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1587 -----LSEEVVPEELVEEV-----VPE-----GLVE-EVVPPEELVFLPEELVEE 1625
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1385 VIETVVISETGES-----PECVGAHLLPAKSSATGHHWTLOHAEDTVPGLGESQAESI 1438
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1626 VPEVVV-EEVEEELVBEFIPPEEVEEVLVPEE-----LVPEKLTIEDV 1665
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1439 PIIVTAPESTLHPDLOGRISASQRSERSEEDKPDAGDADGKSTAEIKVLKAEPEILE 1498
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1666 -----VPEEIVEEVLPEEV---VKEVPEEVEEVLPEELVEEVEEVEEVEEVIPEELI 1716
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1499 LESKSNKIVLNVITQAV 1515
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1717 AEVVSQLEVEEVLPEEV 1733
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 13
A56577
microtubule-associated protein MAP 1B - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C>Date: 21-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 16-Feb-1997
C:Accession: A56577
R:Zauner, W.; Kratz, J.; Staunton, J.; Feick, P.; Wiche, G.
Eur. J. Cell Biol. 57, 66-74, 1992
A:Title: Identification of two distinct microtubule binding domains on recombinant rat
A:Reference number: A56577; PMID:92347374; PMID:1639092
A:Accession: A56577
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-2364 <ZAU>
A:Cross-references: GB:X60550
A:Experimental source: brain
A>Note: nucleotide sequence not given; conceptual translation not complete
C:Superfamily: microtubule-associated protein MAP1B

Query Match 5.8%; Score 470; DB 2: Length 2364;
Best Local Similarity 20.2%; Pred. No. 1.9e-09;
Matches 328; Conservative 234; Mismatches 602; Indels 458; Gaps 62;

QY 45 ADPATKL-----PQKQQLSSVNGVAEQGDVH-----VOEENQEQEVEEVDVQGRE 93
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 382 ANPAEKIIRVLPFGNSTQNYLSEKLEKHLDFLKQPLATQKLTGQVSTPPVKQVKLKQ 441
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 94 SEDVRE--KORVEEMAANSTAVEDITKQGEETSEIEIQIPASENNVVEEVMQPAESQAND 151
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 442 RADRSRESLKPATPLPSKSKS-----VRKESKEAPE-----ATKASQVEKTPKVESKE-- 488
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 152 VSGFKKVPKFGVGFYFTKDKNEKSDTVQLLTVKKDEGEAGASVAGADHQPSPVETAVGE 211
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 489 -----KVIVKKDKPGKVESKPSVT-----EKEVPKSEQSP-VKAEVAE 526
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 212 SASKESELKQSTEBQGTUKQEOSTEIPLOAESDQAAEEAKDEGEKEEPTKSPES 271
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```

Db 527 KRAATSKKRVTKD---VVKKE-----IKTKPEKKEEKKKAVAKKEDTPLKKDEK 576
Qy 272 PSSPVNSETTSFKKFTTHGAWGRKRTSPFKSKEDDLTAARKEKQEAKEYDEEEK--E 329
Db 577 PK---KEEAKKEIKK-----EIKKEEKKELK-KEYKKEETPLKDAKKEVKOE 619
Qy 330 KTEPASEOEPEADTDQAALSDYKEVELPLEDQVODDLASSEKCAPLATEVPEDEKMA 389
Db 620 KKEVKEEKEPKKEI---KKISKDIKK-STPLSD-----TKKPAALKPKYAKKEEPT 667
Qy 390 HOEVAEAVSVTEKTEEEOGGGEAGGVVETGESLPPREKLAEPQAEVPOEAPEBEL 449
Db 668 KKEPILAGLKKKGGKAVTKKGGKTEAATAVGTAAVAAAAGV-----ASGAKEL 720
Qy 450 MKSREMGVSGDHTQ-LTDLSPBEKTLPPHPEGIVSEVMSSQERIKVQSGPLKKLFSS 508
Db 721 BAKRSIMSSPEDLTKDFEELKAEIDVAK---DIKQLELIEDEEKLK-----765
Qy 509 SGLKKLSKKQKQKGGGGDEBERGEYQHTHSPESADQKGBSSASSPEBPEETCLEK 568
Db 766 -----ETEPGEAYVIQKE-----TEVSK 783
Qy 569 GPLBAPOGEAEGETTSDEKKREBITTPWASFKQWTPKKAVRPSPESDKEELEKVKSA 628
Db 784 GSABSPDEGI---TTBEGEGE-----QTP-----EBLEPV---812
Qy 629 TLSSTDSTVSEMODEVKTVEBQKPEPRKRVDTSVSWMALICVSSKKRARKASSDDE 688
Db 813 -----EKQGVDDIKFEDEGAGFEESBAGDYEBKA-----843
Qy 689 GGPRTLGSHAEFASKKEKAGTAVPASTQEQDQACSSSPERPGSSBEGVSTWES 748
Db 844 -----ETEEKEEPEEDG-----EDNVSSASAH---SPTDEBIKAKA 879
Qy 749 FKRLVTPRKSKSLKEKEDSSVEQLST-EIEPSREESWVSIKIPRRKKRADGKOE 807
Db 880 DVH-IKEKESVASGDDRAEBMDLEKGEAQAQSEEG-----EEEDKAE 925
Qy 808 QATVEDSGVEINEDDPNVPVAVPSEYNAVEREKMAQCNTELPGLLGAVVSEELSKT 867
Db 926 DAREEDHEPDTEABD-YMAVAVDKAAGVTEPDYDFLGT---PAKQGVQSPSREPAS 981
Qy 868 LVHTVSAVIDGTRAVTSVEER---SPSISAS--VTPELEHTAGAMPVEVEYFKDII 922
Db 982 S1HDETLPGSSSEATASDENREDDPEEFTATSGYOSTIEBSSPTPMDEKSTPRDWM 1041
Qy 923 ABEET-----PVLT--QTLPEKG--DAHDMVTSEVD 949
Db 1042 TDETNNEETESPQEFVNITKYESLSYQEFYSKFPVASFNGLSGSKTDATDGR-----D 1096
Qy 950 FTSEAVTATETS-----EALRTEVTEASGAETTDWVS-----AVSOLT 989
Db 1097 YNASSSTISPPSMEEDKFSKALRDAPPEETDVKTAGELDKOVSDRBLSPAKSPSLS 1156
Qy 990 DSDPTTEATPVQOEVESGLTDEBERQTOALLQAVADKYKESQVPAQTOYQRTSKAL 1049
Db 1157 PSPSPETIKTPGERSVNSLTPE-----IKASABEBAVAVSPVTOAV 1202
Qy 1050 -----EKVEVEDESEVLASEKEKOVMPKPVQOAGAEHLAAGSETGOATPESLEV 1100
Db 1203 VEEHCASPEEKTLEVVSPQSVTGAAGHTPYQSPTEKSSSHPLEVTEHQAQVAPVSFEF 1262
Qy 1101 PEVTADVHVATC---QYIKLQOLMEQAVAPESSETLTDSETNGSTPLADSPYADGTQ 1156
Db 1263 TEAKDENERSSTSPMDEPVPSDESPIEKVLSPLRSPPLGSEBAVDFLSADDKALGRS 1322
Qy 1157 DETIISQDSKATAAAYQOVTREBAATAQKEBPTLPNNVPAQEEGEEBGRDYLEPTQ 1216
Db 1323 ESPFGKNGKQGFSPDKESPVSDLTSLYODKQEKRAKAGFIPIKEDSPPKKASDAIMSS 1382
Qy 1217 ELTAAVAVPLAKTEVQOEG---EVDMLDGEKYKEOEVEFVHSGPNQKA-----ADVT 1266

```

```

Db 1383 QSALA-----LBERKLGDGSPPTQYDVQSGSFKEDETKMSISEGTVSDSKATPVDEGAEDT 1438
Qy 1267 YDEVMVAGCQ-----EKSTEV-QSLSEBGE-----1294
Db 1439 Y-SHMEGVASVTSVATSSPPEPTDDVSPSLHAEVSGSPHSTEVDLSVSVQTPPTF 1497
Qy 1295 METDVEKERE-----TKREQVSEBGEQETAPEHGHTGKVPULTLDMSSSEKGLG 1347
Db 1498 QETEMSPSEKEPRPMSISPPSPFKTAKSRTPVODHRSBQSSMSIERFGQSPERSHAMD 1557
Qy 1348 SLGSPSLPDQKAGCI-----EVQVQSLDPTTVQTAANV-----EKVIET 1388
Db 1558 FSRQSPDHPTVG-AGMHITENGPTEVDYSPSDIDQSSLXHKIPTTEPSYTOQNDLSEL 1616
Qy 1389 VVISTGSESPCEVGAHLLPAKSSATGSHWTLQHAEDTVPKPSQASSTIPIYTPAPES 1448
Db 1617 ISVSQVEASPTSSAH-TPSQIAS-----PLQEDTLDSD---VPPPRDM 1655
Qy 1449 TLHPDLQGE-ISAQGR-----SEBEDKPDAGR-----DADKESTALEKVLKA 1492
Db 1656 SLVYASLASEKVQSLGEKLSPKSDISPLTPRESSFTYSPGFSSTGAKESTAYOTSSS 1715
Qy 1493 EP 1494
Db 1716 PP 1717

```

```

RESULT 14
D96796
probable heat shock protein, 53413-59028 [imported] - Arabidopsis thaliana
CSpecies: Arabidopsis thaliana (mouse-ear cress)
CDates: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
CAccession: D96796
Rthnology: A.; Becker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huzar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Matzli,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: D96796
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1871 <STO>
A:Cross-references: GB:A805173; NID:96143906; PIDN:AAFO4452.1; GSPDB:GN00141
C:Genetics:
A:Gene: F28016.15
A:Map position: 1

```

```

Query Match 5.8%; Score 466.5; DB 2; Length 1871;
Best local similarity 20.5%; Pred. No. 1.9e-09;
Matches 331; Conservative 248; Mismatches 588; Indels 447; Gaps 68;

Qy 50 KLPQKNGQLSSVNGVAGQDVHVOENQ-----GQEEVVED 88
Db 310 KLOKEBQHSK---EQGG--HSKEENKELVEEKTPEAETTRINDILGQGEIIVPEVD 363
Qy 89 VQGRSEDEVREKDR-VEEMANSTAVEDI-TKQGEELTSLEIQIPAE-----N 136
Db 364 TLGKTSDEGKEKQNTVKEIKNGATEIEIDKMGVFASNIADTGMNSEDPSDKLESAD 423
Qy 137 NVEEVPQAE-----SOANDVGFKRVFKFTVKKDKNKKSDTVQLTLTKKDE 187
Db 424 EYDKAVVEKKDROEENDKXGAQSEDSLTKLQEIGGQYQGGK-RIDKQENIKELR---E 478
Qy 188 GEGASAVGAGDHQBPVSVEVAVGESAKSELSKOSTEK--OEGTLKQBSSTELPQAS 245
Db 479 GQASAA-----EKNIKNDILKRVQKRSBGKHKIQTPEETNKQEGVNEKIM--ET 528

```

```
QY 246 DQAAEEAKDEGEE---KOE-KEPTKSPSPSP--VNSSETSSFKFFTHGWGHRKKT 299
Db 529 GKINEDGTRKVOEMIRQOELDPASEKENRSRELVKNTNDEKK-----EKBI 579
QY 300 SFKSKEDDLTAETAEKKEQB-ABKVDE-----EKEKTPASEEOEPAEDTDQARL 349
Db 580 ACTERKESDRPKILREQEVADEVAEDTKPSIYGEVKEEBIAGKEKEFGSDDDIARI 639
QY 350 SADYEV-----ELPLEDQVGD-----LEASSEKCAPLATEVDFDEK 386
Db 640 VRDTEQLDSNAMQOBEKDMIQELVLEEKVDCGGKGIILVAETKAENNKSKRVOETEEOQ 699
QY 387 ME-----AHQEVVAHVHSTVEKTEEEQCGGGAEGGVVEGTGSLPPEKLAEPQEV 440
Db 700 LDKCDGKHFOKLIIEGISHDGEVDEKGGKRTA-----EKRIK 741
QY 441 QEAPEAEELMKSMCMVSGG--DHTOLTDL-----SPEKTLPKHPGIV----- 483
Db 742 DRAREAEII-KEKDLGVSGRYIKGTWIKELVENRGYIRNEHEEKKDDANRPEKITGTIK 800
QY 484 SEVEMLSQ-----ERIK-----VOGSPLKILFSSSGLKKLSGKKQ--- 519
Db 801 QELVLSNQLROBNVEDGDKTQELVEEKIKDCEEBEGSESKITDDVVRKVQGIKEEEL 860
QY 520 -KKGKGGGDEEPCGYQHHTHTEPSADEQKGSASSPEEPETTCLEKGPLAEPODGE 578
Db 861 YKPKRHHG-----TKTILVEETG--YEQKEKET-----AESDIE 896
QY 579 AEGGTT--DGEKKREGITPWSFKMWTPKRVRRPSESDEKEELEK-VKSATLSTDS 635
Db 897 AECGSLRKVDGIEHELHPKIHKEKRNRRVTGAKPSGOEKEKEKIVESMTTENDN 956
QY 636 TVSEMDEKVTG-----EOKPEPKERVDTSVSWEALICVGSS 675
Db 957 SIDVQETKXKPCRLSHDKRYKIQELLMAGHNDKKEEQNENVAEVELET---ERVS 1013
QY 676 KGRARKASSDDGG-----EOKPEPKERVDTSVSWEALICVGSS 701
Db 1014 SKKVQSGMEDDNGSGFHFEERKSYEDWTHKEKREKVLVBEETYPKDKHTGGEDHND 1073
QY 702 EBASKDKEAGTDAVPASTQ-----ODAQSSSPPEAGSPSEGEVSTWESFK 750
Db 1074 HKEEBQKENVIAKAEINTEDESPKVEETEKQDHGELKRSWQAKKQETEEKDKTRAMEK 1133
QY 751 RLVTPRKK-----SKSL-----EKAEDSSVEOL-STEIEPSRE-----ESW 787
Db 1134 NETVRRKQTKDGLKREGEPELGGHERRGEEDRIEVLVETISDHKEVKYKKDEY 1193
QY 788 V-----SIKFIPIGRKKRADKQEQATVEDSGPVEINEDDPNVPVAVPLSE-----YNA 837
Db 1194 ILRSQDTGKVDLGERERRSKRKIKHSVEDEIGDQEDAEAEAAVVRNENGSSRKVQT 1253
QY 838 VEREKMAQNTLPGLLGAVVYSELSKTLVHTVSVAVIDGTRAVTSVEERSPWSIAS 897
Db 1254 IEESKHKQNKIPETSNP-EVNEE-----DEERV-----EKTKEVEAH 1294
QY 898 V-----TEPLETAGAMPVPEVTEKDIABETPVLTOTLPEGKADHDMVTSEVDFT 951
Db 1295 VQLEGKTENCKDDGEGR--BERGKQGTAE--MLRQR---KTKSDDGIVRKIQT 1347
QY 952 SEAVTATETSEALRTEVTASGABETDMVASVQLTSPDPTTEBATPVQEVESGLDT 1011
Db 1348 KE-----BEPDEKKSQESSSHVVKLVAE--DG-----SLRNG-LEF 1380
QY 1012 EEEERQTOALQOAVDKVKEESQVPATQTVORTGSKALEK-----VEEVEDSEVLAS 1064
Db 1381 SEKESTVSKMLKLDSEKKEHKIRKPTTEERSNAPVIEKQGNKNAEEMQDKIDRRGK 1440
QY 1065 EKE-KDVMKPGPVQEAAGHAQ--GSETGQATPESLEVPVTDVHVATQCVIKIQ 1120
Db 1441 NQBIKQEPYGVLRNHEHDKITETHRGEKGT-----ENSS---TKIQ 1483
QY 1121 LMEQAVAPESSETLTDSETNGSTPLADSDTADGTQOQDETIDSDSKATAAVRSQVTEE 1180
```

```
Db 1484 TQDELEKPKRPSISENNN-----IHFMDSSQDIEEKGSQOAEKYAKQNKIQE--- 1534
QY 1181 AATAQKEEPTLPPNNVPAQEEHGEPRDVLPTQOELTAAAVPVLAKTEVQO-----E 1234
Db 1535 -----VMNDEKKEEY-----HISERVNEMAKRILQVESKANDGSSKKNETE 1577
QY 1235 G-EVDWLDEKVKYEQEVFVHSGPNQSKAADVTYDSEVNGVAGCQKKESTEVQSLSLEG 1293
Db 1578 QGESTGLRKRKREHQLVELETSDQKKG--VKDEVVGKAEIIEDEYDSSRKIHEHEE 1635
QY 1294 EMETDVEKRETKPEQVSEGEQETAAPHEHGTGKPVLTLDMPSSERGGKALGSL--- 1349
Db 1636 RMSDKLMEHGEEMBEKJAEETSDEGAENGRAG-----KSRDDGFGKVRKIE 1686
QY 1350 ----GGSPLPDQDKAGCIEVQVQSLDTTIVTQTAABAVEKVIETVVISSETGESPE 1399
Db 1687 VQKNDQSFVEKDTSGRAKENLNDDEPTKTETKATDNESRKIHQIKEOGTSEQ 1740

RESULT 15
T34518
nestin - golden hamster (fragment)
C:Species: Mesocricetus auratus (golden hamster)
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T34518
R:Steinert, P.M.; Chou, Y.H.; Prahlad, V.; Parry, D.A.D.; Marekov, L.; Wu, K.C.; Jiang, S
submitted to the EMBL Data Library, December 1998
A:Description: A high molecular weight intermediate filament associated protein in BHK-2
with type III vimentin and type IV a-internexin.
A:Reference number: Z21538
A:Accession: T34518
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1804 <STE>
A:Cross-references: EMBL:AF110498; NID:g4063502; PID:g4063503; PIDN:AAC98312.1
A:Experimental source: cell line BHK-21

Query Match 5.7%; Score 456.5; DB 2; Length 1804;
Best Local Similarity 20.8%; Pred. No. 4.2e-09;
Matches 358; Conservative 249; Mismatches 575; Indels 541; Gaps 79;

QY 2 GAGSSTEQRSPPQ-----AGSDTPELVLVSGHPAAEASGAAGDADPADATKLPQKN 55
Db 246 GVLTELEAGGQPGHPEDATASAPS---LSPHPVLEAKD--GDSTES----- 290
QY 56 GQLSSVNGAEOGDVHVQENQOGEVEEVVDVQGRESEDEVREKDRVVEEMAANSTAVED 115
Db 291 -----RGSSIFQED--EGQIWELV-----EKEAAIELKVESSLAQE 324
QY 116 ITKDQOETSETIEQTPASENNVEMVOPAESQANDVGFKVFKFVGFKFTVKDKNEKS 175
Db 325 TQEDGLH-----TEIQQDSQGLQK----- 344
QY 176 DTVQLLVTKDGEGBAASVAGDHOEPSEVETAVG-ESAKSESELKQSTKEGTLKQBQ 234
Db 345 ETLEAL-----GESPLSLKIQNHETPGKNCNLSRSDENQGTLLKSPEEEKQTLKSL 398
QY 235 SSTEPLQAESDQAAEEAAKDEGEKQEKPTKSPSPSPVNSSETTSSFFKFFTHGWAG 294
Db 399 BEKDQVEVEKTLKGVPELSKPLGKE---DPRIEQELMSPEGTLETLSF-----IG 446
QY 295 WKTKTSFKKSDEDDLET--AEKRKEQAEKVDEEKEKTEPASEOEPAEDTDQARLSAD 352
Db 447 KNEEVVRSSEENETESLAAPFKESQHPGLGCEEBEIORVERLIEKE-----QOESLSP 500
QY 353 YEKVELPLEDQVGD--LEASSEKCAPLATEVDFDEKMEAHQEVVAHVHSTVEKTEEBEQ 410
Db 501 EE-----EDQETDRPLEKENGEPKLPVEE--DQIFETLIEKEGQESLSSPEEEDQET 552
QY 411 GGEBAEGVVVGTGESLPPPEKLAEPQEVQAE-----PABELMKSRMCMVS 458
Db 553 RPLEKEEDQLVERLVKEGQESLSSPEEEDQETDRPLEKENGEPKLPVEEEDQLFETLIE 612
```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2002, 21:09:06 ; Search time 49 Seconds
(without alignments)
1350.943 Million cell updates/sec

Title: US-09-902-432-4

Perfect score: 8073

Sequence: 1 MGAGSTQRSPQAGSDT.....AWAQRKCLPRLQLKAPVSK 1596

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4294	53.2	1781	1 AKAC HUMAN	Q02952 homo sapien
2	501	6.2	2464	1 MABE MOUSE	P14873 mus musculus
3	493	6.1	3924	1 ANK2 HUMAN	Q01484 homo sapien
4	492	6.1	2468	1 MABE HUMAN	P46821 homo sapien
5	470	5.8	2459	1 MABE RAT	P15205 rattus norv
6	433.5	5.4	2774	1 MAPA RAT	P34926 rattus norv
7	421	5.2	2805	1 MAPA HUMAN	P78559 homo sapien
8	411.5	5.1	3562	1 PGCV CHICK	Q90953 gallus gall
9	408.5	5.1	4687	1 PLEI RAT	P30427 rattus norv
10	402.5	5.0	4473	1 PLEI CRIGR	Q91555 cricetus
11	400.5	5.0	4684	1 PLEI HUMAN	Q15149 homo sapien
12	395	4.9	3396	1 PGCV HUMAN	P13611 homo sapien
13	388	4.8	1805	1 NEST RAT	P21263 rattus norv
14	388	4.8	1861	1 MAP2 RAT	P15146 rattus norv
15	385.5	4.8	1087	1 NPH MOUSE	P19246 mus musculus
16	382	4.7	2116	1 MYS2 DICDI	P08799 dictyosteli
17	379	4.7	2349	1 TPR HUMAN	P12270 homo sapien
18	377.5	4.7	1827	1 MAP2 HUMAN	P11137 homo sapien
19	376	4.7	3256	1 KI67 HUMAN	P46013 homo sapien
20	374	4.6	1233	1 YF16 YEAST	P43597 saccharomyc
21	374	4.6	1637	1 MRSP STRAU	P80544 staphylococ
22	371.5	4.6	1828	1 MAP2 MOUSE	P20357 mus musculus
23	371.5	4.6	3911	1 AKA9 HUMAN	Q99996 h a-kinase
24	371	4.6	1658	1 M67 YEAST	Q03661 saccharomyc
25	370	4.6	1391	1 MS72 DROHY	Q08696 drosophila
26	369.5	4.6	3381	1 PGCV BOVIN	P81282 bos taurus
27	367.5	4.6	1411	1 TCOF HUMAN	Q13428 homo sapien
28	367	4.5	1020	1 NPH HUMAN	P12036 homo sapien
29	366.5	4.5	1972	1 P531 HUMAN	Q12888 homo sapien
30	363	4.5	831	1 NPH RAT	P16884 rattus norv
31	363	4.5	2453	1 NCRI MOUSE	Q60974 mus musculus
32	354	4.4	1618	1 NEST HUMAN	P48681 homo sapien
33	352.5	4.4	1616	1 P200 MYCGE	Q49429 mycoplasma

RESULT 1

AKAC_HUMAN

ID AKAC_HUMAN STANDARD; PRT; 1781 AA.

AC Q02952; Q99970; Q00498; Q00310;

DT 01-FEB-1994 (Rel. 28, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE A-kinase anchor protein 12 (A-kinase anchor protein 250 kDa) (AKAP

DE 250) (Myasthenia gravis autoantigen gravin).

GN AKAP12 OR AKAP250.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A. (ISOFORM 1).

RC TISSUE=Heart;

RX MEDLINE=97153077; PubMed=9000000;

RA Nauert J.B., Klauk T.M., Langeberg L.K., Scott J.D.;

RT patients, is a kinase scaffold protein.";

RL Curr. Biol. 7:52-62(1997).

RN [2]

RP SEQUENCE FROM N.A. (ISOFORM 2).

RC TISSUE=Umbilical vein endothelial cells;

RX MEDLINE=98269042; PubMed=9604001;

RA Sato N., Kokame K., Shimokado K., Kato H., Miyata T.;

RT "Changes of gene expression by lysophosphatidylcholine in vascular

endothelial cells: 12 up-regulated distinct genes including 5 cell

growth-related, 3 thrombosis-related, and 4 others.";

RL J. Biochem. 123:1119-1126(1998).

RN [3]

RP SEQUENCE OF 43-1781 FROM N.A.

RC TISSUE=Umbilical vein endothelial cells;

RA Bowditch R.D., Ginsberg M.H.;

RT "Sequence of gravin cDNA isolated from a human umbilical vein

endothelial cell library.";

Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.

RL [4]

RP SEQUENCE OF 1477-1781 FROM N.A.

RC TISSUE=Umbilical vein endothelial cells;

RX MEDLINE=92395179; PubMed=1522245;

RA Gordon T., Grove B., Loftus J.C., O'Toole T., McMillan R.,

Lindstrom J., Ginsberg M.H.;

RT "Molecular cloning and preliminary characterization of a novel

cytoplasmic antigen recognized by myasthenia gravis sera.";

J. Clin. Invest. 90:992-999(1992).

RL [5]

CC FUNCTION: ANCHORING PROTEIN THAT MEDIATES THE SUBCELLULAR

COMPARTMENTATION OF PROTEIN KINASE (PKA) AND PROTEIN KINASE C

CC (PKC).

CC SUBUNIT: BINDS TO DIMERIC RII-ALPHA REGULATORY SUBUNIT OF PKA.

CC SUBCELLULAR LOCATION: CYTOPLASMIC. MAY BE PART OF THE CORTICAL

CYTOSKELETON.

CC ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; MAY BE

PRODUCED BY ALTERNATIVE SPLICING.

CC TISSUE SPECIFICITY: EXPRESSED IN ENDOTHELIAL CELLS, CULTURED


```

CC FIBROBLASTS AND OSTEOSARCOMA BUT NOT IN PLATELETS, LEUCOCYTES,
CC MONOCYTIC CELL LINES OR PERIPHERAL BLOOD CELLS.
CC -1- INDUCTION: ACTIVATED BY LYSOPHOSPHATIDYLCHOLINE (LYSOPC).
CC -1- DOMAIN: POLYBASIC REGIONS LOCATED BETWEEN RESIDUES 265 AND 556 ARE
CC INVOLVED IN BINDING PKC.
CC -1- DISEASE: ANTIBODIES TO THE C-TERMINAL OF GRAVIN CAN BE PRODUCED BY
CC PATIENTS WITH MYASTHENIA GRAVIS (MG).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation --
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb.ch/announce/
CC or send an email to license@isb.ch).
CC -----
CC
CC DR EMBL, U81607; AAC51366.1; -.
CC DR EMBL, AF001504; AAB58938.1; -.
CC DR EMBL, AB0013476; AAB19927.1; -.
CC DR PIR, M96322; AAA35931.1; -.
CC DR PIR, A43922; A43922.
CC DR GeneW: HGNC:370; AKAP12.
CC MIM: 604698; -.
CC
CC DR InterPro: IPR001573; PkinA anch.
CC DR Antigen; Repeat; Alternative splicing.
CC KM DOMAIN 603 633 AKAP 1.
CC FT DOMAIN 752 782 AKAP 2.
CC FT DOMAIN 797 827 AKAP 3.
CC FT DOMAIN 98 101 POLY-GLU.
CC FT DOMAIN 265 556 INVOLVED IN PKC-BINDING (PROBABLE).
CC FT DOMAIN 1540 1553 RII-BINDING (PROBABLE).
CC FT VASPLC 1 98 MISSING (IN ISOFORM 2).
CC FT VASPLC 99 106 EEEVYTE -> MGLTITIT (IN ISOFORM 2).
CC FT CONFLICT 117 117 E -> K (IN REF. 2).
CC FT CONFLICT 142 144 KKN -> TPEI (IN REF. 2 AND 3).
CC FT CONFLICT 215 215 G -> K (IN REF. 2).
CC FT CONFLICT 448 448 G -> E (IN REF. 2 AND 3).
CC FT CONFLICT 694 694 R -> G (IN REF. 2 AND 3).
CC FT CONFLICT 867 867 G -> S (IN REF. 2 AND 3).
CC FT CONFLICT 986 986 S -> A (IN REF. 2 AND 3).
CC FT CONFLICT 1330 1330 S -> EE (IN REF. 3 AND 4).
CC FT CONFLICT 1581 1581 V -> M (IN REF. 4).
CC FT CONFLICT 1601 1601 Q -> L (IN REF. 2).
CC SEQUENCE 1781 AA; 191439 MW; BA813937379FAC0F CRC64;

```

Query Match	53.2%;	Score 4294;	DB 1;	Length 1781;
Best Local Similarity	57.9%;	Pred. No. 7.8e-140;		
Matches 959;	Conservative 187;	Mismatches 401;	Indels 108;	Gaps 35;

Oy	1	MGAGSTQRBPEQ	PAGSDTPSELVLSHGHPALFAS	-GAAGPADA	-DPATLPOKNG	56
Db	1	MGAGSTQRBPEQ	PRESSSTPAEBEPGGPSAEAPD	TADPALIASPATLLOKNG		60
Oy	57	OLSSVNGAEOGDVHVOENQGO	-----EEEVDEDVGOSESEVREKDR	1022		
Db	61	OLSTINGAEODELSLQGDGLNGQALNGQALNSQOESEEVITVEYQGRDSEVSRDS	120			
Oy	103	VEEMANSTAVEDITLKDOQOEETSEIIEQIPASENNVEEMVPAESQANDVGFKKVEFVG	162			
Db	121	DKEMATKSAVHIDITDDQOEENRN-IBIPSPESNLEITLOPTESQANDIGFKVFEFVG	179			
Oy	163	FKFTVKDKNEKSDTVOLLTVKKQDEGEGAEASVAGDGHQEBSEVETAVGESAKESKELKOS	222			
Db	180	FKFTVKDKTEKPTVOLLTVKKQDEGEGAEASVAGDGHQEBSEVETAVGESAKESKELKOS	234			
Oy	223	TEKQEGITLKQBOSSTEIRPLQABSDQAAEEKAKDGEKQEKAPTKSPESPSPPNNSSTIS	282			
Db	235	TEKPEETLKREOSHAEISPPAPESQAV-ECKEKEGEKEQEKEPKSAESPSTSPSTSETGS	293			
Oy	283	SFKKFTGHWAGMKRKSTSPKSKSDEDDLETAKRKQEKQEKVDEBEKKTEPBASE	337			
Db	294	TEKKFTGHWAGMKRKSTSPKSKSDEDDLETAKRKQEKQEKVDEBEKKTEPBASE	353			

QY	338	---	OEBAEPTDQARLSIXDEKXELPLEEOVODLEASSSEKCAPLATVPEPKKBAH-O	E	392
Db	354	QAHPOEPESHAHERISAETYEKVELSEBQVSGSQSPSEKPAPLATVPEDEKILEVOEB	E	413	
QY	393	VVAEVHVSATVEKTEEBEEOGGGEMAEVGEVGESLPEPKLAPOEVEQOEAPEAEELMS	452		
Db	414	VVAEVHVSATVEBERTEBQ-----KTEVEBFRAGVPAEELVGMAREQOEAPEAKELVKL	465		
QY	453	REMCVSGDHTQULTDSPEBKTLPKHPEGIVSEBVELSSOEERIKVQSGPLKCLPSSGLK	512		
Db	466	KETCVSGEDPFGQADLSPDEKVLSPPEGVSEVEMLSSOEEMKVQSGPLKCLFTSTGLK	525		
QY	513	KLSEKKQKRGKGCGGDBEPPEYHITHSPESADQOKESASSSPSEBEBETCLEKEGLE	572		
Db	526	KLSEKKQKGR-KGGGESESEHQVPADSPDQOEKSGSSSPSEBEBETCLEKEGLAE	584		
QY	573	APQOGEAEBGTSNGEKKRGITPMASFKKMTPKRVRPSESPKEELEKVMATLSS	632		
Db	585	VQOQGEAEBGATSDGEKKRGVTPMASFKKMTPKRVRPSESKEDBLDKVMSATLSS	644		
QY	632	TDSTVSMEDEVKTVEEOKPEEPKRRVDTSVSWEALICVSGSKKRRARKAASSDDBGPR	692		
Db	645	TESTASMEQOEMKOSVEPEPPEPKKRVDTSVSWEALICVSGSKKRRARRSSDEBGPX	704		
QY	693	TLGGDSHRAEASXDKACAGTDAVPASTQOQOQAQSSSPPEPAGSSBEGVSTWESFKRL	752		
Db	705	AMGDBHQKADBAKGDKXTGTGDILASQEHDPQOGSSSPEQAGSPTBEGVSTWESFKRL	764		
QY	753	VTPPKSKSKLKEEKAED---SSVEQLSTEIPEPSRESVNSIKKTIPOKRRKPADGQEQ	808		
Db	765	VTPPKSKSKLKEESSEDSINGSGVENHSTPDEPKGEESVNSIKKTIPOKRRKPADGQEQ	824		
QY	809	ATVEDSGPEINEDPNVPAPVAVLSEYNAREKME---AQNTLEPOLLAGVAVVSEELS	865		
Db	825	APVEDAGPTGANEDDSPVPAPVAVLSEYDAVEREKMEQAQAGAPQQAQKATEVSKELS	884		
QY	866	KTLVHVSVAVIDGTRAVTSVEERSPSMISASVTEPLEHTAGEAMPVPEVTEKDIIA-E	924		
Db	885	ESQVHMAAAVADGTRATIIIEERSPSMISASVTEPLEHVEAEMALLTEVLEVERIEAE	944		
QY	925	ETPVLTOTLPEPKOAHDMWTSEVDPTSEAVTATEPSEALRTBEVTAESGAETTDWISA	984		
Db	945	EPPVTEBPLEPNKRAKDDIVSSAEITPEAVTAEAGLSEBGETESAABETTEMVSA	1004		
QY	985	VSOLTSDPTTEATPVQOEVESGVLDTBEERQOALLOAVADKVKESQVPAT-----	1038		
Db	1005	VSQLTSDPTTBEBATPVQOEVGGVPDIEQERTEQVLOAVAEKXESQLPGTGGPEDV	1064		
QY	1039	-QTVQRTGSKALEKVEVEEDSEVLASEKRDVMPKPVQOEAHEHLAQSGETQATPES	1097		
Db	1065	LQPVQRA-----EAERPEQOAEASGLKKEITDVLVDAQOEAKTPEPTQKGVAGGTPTES	1118		
QY	1098	LE-VPEVADVDH---VATQOV-----IKLOOL-MEONAPAPSSSTLUDSENGSTPLAD	1147		
Db	1119	FEKAPQVTESESEELVTTQOAEFTLAGVNSQOEWEQALPRVSVETPTDSETDGTPPAD	1178		
QY	1148	SDTADGTQODETTISODSKATAVROSQVTEEPAATAQKEBPSTLPNNVPAQOEHGEEBP	1207		
Db	1179	FDACGTTQOKEIYIEHNEENVASGTOSGGTEAEVAPQAKERPAP-PSGFVPOEETKEBKS	1237		
QY	1260	QKADVTYDSVMKAVAGCQKEKSTEVOS-----LSLEBEMETDVEKKEKTEKPEQVSE	1314		
Db	1296	EKVTEVALKGTGEAEACKDQDALEQSAKSPSPSEVEEMVAVQERKTEAEPTHVABE	1355		
QY	1315	G-BOETAPABHEGTVGKPVVTLTMPSEBKGALGSI-GGS-PGLPQODRAGCLEVVOYSLD	1372		
Db	1366	KLEHETAVTVSEBVSQOLDTQVAVPILIDAKEVYSLEGSPPCLQOEBEAVCTKIVQOSE	1415		
QY	1373	TTVQVTEAENAEKVI--ETVVISSTGSPRECVGAMHLLPAKSSATGGMHTLQHAEDTVLGG	1430		

```

Db 1416 ASFTLTAAREEKVLGTANILETGETLEPAGAHVLEESSEKNEDFAHPGEDAVPTG 1475
QY 1431 PESQAESPIIVTAPPESTLHPDQGISASQERSEEDKPDAGPDADGKSTAJEKVL 1490
Db 1476 PDCQAKSTPVIIVATTKGLSSDLEGEKTTSLKWKSEDEVQVACQEV--KVSVAIEDL- 1532
QY 1491 KAPEE--ILELESKNKVLNLTQVADQPART-ETAPETHAYDSQTVQVACLDSEPN 1547
Db 1533 --EPENGILSETKSKLVNIIQTVADQVQTEETATEMLTSELOTQAHVIKADSQDAG 1590
QY 1548 RCWTK-----MKDAKMKHPVPQPREDLQVLTVLEA 1577
Db 1591 QTEKESGEEFQASQADETPITSKESESTAVGQA 1625

RESULT 2
MAPB MOUSE
ID MAPB MOUSE STANDARD; PRT; 2464 AA.
AC P14873;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Microtubule-associated protein 1B (MAP 1B) (MAP1.2) (MAP1(X))
DE [Contains: MAP1 light chain LC1].
GN MAP1B OR MTAP1B OR MTAP5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., AND DOMAIN.
RC STRAIN=Swiss Webster; TISSUE=Brain;
RX MEDLINE=90094539; PubMed=2480963;
RA Noble M., Lewis S.A., Cowan N.J.;
RT "The microtubule binding domain of microtubule-associated protein
RT MAP1B contains a repeated sequence motif unrelated to that of MAP2
RT and tau."
RL J. Cell Biol. 109:3367-3376(1989).
CC -!- FUNCTION: THE FUNCTION OF BRAIN MAPS IS ESSENTIALLY UNKNOWN.
CC PHOSPHORYLATED MAP1B MAY PLAY A ROLE IN THE CYTOSKELETAL CHANGES
CC THAT ACCOMPANY NEURITE EXTENSION. POSSIBLY MAP1B BINDS TO AT LEAST
CC TWO TUBULIN SUBUNITS IN THE POLYMER, AND THIS BRIDGING OF SUBUNITS
CC MIGHT BE INVOLVED IN NUCLEATING MICROTUBULE POLYMERIZATION AND IN
CC STABILIZING MICROTUBULES.
CC -!- SUBUNIT: 3 DIFFERENT LIGHT CHAINS, LC1, LC2 AND LC3, CAN ASSOCIATE
CC WITH MAP1A AND MAP1B PROTEINS.
CC -!- DOMAIN: Has a highly basic region with many copies of the sequence
CC KKEE and KKEI/V, repeated but not at fixed intervals, which is
CC responsible for the binding of MAP1B to microtubules.
CC -!- PTM: LC1 IS COEXPRESSED WITH MAP1B. IT IS A POLYPEPTIDE GENERATED
CC FROM MAP1B BY PROTEOLYTIC PROCESSING. IT IS FREE TO ASSOCIATE WITH
CC BOTH MAP1A AND MAP1B. IT INTERACTS WITH THE AMINO-TERMINAL REGION
CC OF MAP1B.
CC -!- SIMILARITY: TO MAP1A.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X51396; CAA35761.1; --
CC PIR; S07549; QRMSP1.
CC MGD; MGI:1306778; Mtap1b.
CC InterPro; IPR000102; MAP1B_neuraxin.
CC Pfam; PF00414; MAP1B_neuraxin; 10.
CC PROSITE; PS00230; MAP1B_NEURAXIN; 7.
CC Microtubules; Repeat; Phosphorylation.
CC CHAIN ? 2464 MAP1 LIGHT CHAIN LC1.
CC REPEAT 1874 1890 MAP1 1.

```

```

FT REPEAT 1891 1907 MAP1B 2.
FT REPEAT 1908 1924 MAP1B 3.
FT REPEAT 1925 1941 MAP1B 4.
FT REPEAT 1942 1958 MAP1B 5.
FT REPEAT 1959 1975 MAP1B 6.
FT REPEAT 1993 2009 MAP1B 7.
FT REPEAT 2010 2026 MAP1B 8.
FT REPEAT 2027 2043 MAP1B 9.
FT REPEAT 2044 2060 MAP1B 10.
FT DOMAIN 589 787
SQ SEQUENCE 2464 AA; 270408 MW; FBD3DD99CFDBDA87 CRC64;

Query Match 6.2%; Score 501; DB 1; Length 2464;
Best Local Similarity 21.1%; Pred. No. 2.4e-10;
Matches 387; Conservative 266; Mismatches 648; Indels 532; Gaps 86;

QY 45 ADPATKL-----PQKNGQLSSVNGVAEQGVH-----VQENQEQEVEVDEVDVQRE 93
Db 478 ANPAEKIIRVLPFGNSTQYNIILEGLEKHLKLDPLKOPLATQKDLTGQVTPPVKQVKLKQ 537
QY 94 SEDVREKDRVEEMANSTAVEDITDKQGETSEI-----IEQIPASENNVEMVQPAESQ 148
Db 538 RADSR-----LKPATKPVASKSVRKESKEETPEVTKTSQVEKTPKVESKE----- 594
QY 149 ANDVGFKVFKFVGFKFTVKDKNEKSDTVQLLTVKKDEGEAGASVAGDHOEPSVETA 208
Db 585 -----KVLVKDKPKVKTESKPSVT-----EKEVSKSEQSP--VKAE 619
QY 209 VGESASKESELKQSTEQKGTILKQESSTFELPQABSDQAAEEAKDEGEQKEKPTKS 268
Db 620 VAEKQATESKPKVTQDK--VVVKE-----IKTLEEKKEEKPKKEVVKEDKTPKK 669
QY 269 PSP--SSPNSSTTSFKKFFTHGWAGWRKTSFKKSKEDDLTAETAKRKEQAEKVDEE 327
Db 670 DKPRKEEVKKIKKEIKK-----EERKELKVEVKKTPDKAKFVKKEKK--EVK 720
QY 328 KEKTEPASEEQPAEDTDOARLSADYKVELPLEDQVGDLEASSEKCAPLATEVFDE-- 385
Db 721 KEEKEPKKIKKIDIKKSTQSDTKPS--ALKPKVAKKEESTKKE--PLAAGKLKDKG 777
QY 386 -----KWEAHQEVVAEVHVSTVEKTEEEGOGGGEAGGVVVEGTGESLPPEKLAEPQVP 440
Db 778 KVKVIKKEGKTTAAATAVGTAAATTAAVVAAAGIAASGPFVKELEAER---SLMSSPEDLT 834
QY 441 QSAEPAAELMKREMCVSGDHTQLTDLSPREK---TLPKHPEGIVSEVEMLSQERIKV 497
Db 835 KDFEE-----LKAEEIDVAKDIKQLELIEDEEKLKETQCEAYVIQKETEVSKGSAE-- 887
QY 498 QGSPLKLPFSSGLKKGKQKRGGGGDBEPGEYQHIHTESPESADEQKGESSASP 557
Db 888 --SP-----DEGITTTGE-----GECEQTPELEPVEKQGV--DIEKFEDGAGF 930
QY 558 EPEPETTLEKPLAPQDGEAEAGTSDGEKKRGGITPWASFKQKMTVPKRVPRSESD 617
Db 931 EBSSET-----GDYBEKAEETEAEBPEDEGDNASG-----SASK-----HSPTEDD 972
QY 618 KEEEL-----KVKSATLSSTDTVSEMDEVKTVGEEKPEEPKRRVDTSVSWEALICV 672
Db 973 ESAKAADVHLKEKRESVVSGDDRAEEDMDVLEKGEAFQSEEEGEEED----- 1021
QY 673 GSKKRKARAKSSDDEG--GPRTLGGDSHRAEASKDKAGTDAVPASTQEQDQAQSS-- 729
Db 1022 -----KAEDAREEGVEPKTEAEDYVMAVADKAAEAGV-----TEEQYGLGTSK 1067
QY 730 -----SPEPAGS-----PSEGEVSTWESPK-----RLVTPRKSKSKLEKAE 769
Db 1068 QPGIOSPSREPASIHDETLPGGSESEATADENREDQPEEFTATSGVTOGTIETSSRP 1127
QY 770 SSVQELST-----EIPSRSEESVWSIKKF-----IPGR 798
Db 1128 TPMDEMSTPDVMSDETNNETESPSQBFVNITKYESSLYSQEYKPAVASFNGLSEGSK 1187

```

```

QY 799 KGRADGKQOATVEDSGPVEINEDDNPVAVPFLSEYNAREKMEAGQNTLPOLLGAV 858
DB 1188 TDATAQCKDNASASTISPPSSMEEDFKSKA--LRDAYCEEKELASALDIKD----- 1240
QY 859 YVSEELSKTLVHTVAVIDGTRAVTSVEERSPWSIASVTEPLEHT-AGE-----AMP 912
DB 1241 -VSDE-----RLSPA-----KSPS-LSPPSPPIEKTPLGERSVNSLTP 1278
QY 913 VE-EVTEKDIIAETPVLTLQLEKDAHDMDVTSSEVDTSEAVTAT-----E 959
DB 1279 NEIKVAGEEARSVPGLTQAVVEEHCASPEEKTLEVPSPQSVTGSAGHTPYQGPTDE 1338
QY 960 TSEALTE-----EYERAGABERT----- 979
DB 1339 KSSHLPTEVSENAQAVPVSFEFEADENERASLSPMDEBPDSPEVPEKVLSPLSPL 1398
QY 980 -----DMVASVQUL-----TSP-----DTEEATPVQEVES-GVLDTEERROT 1018
DB 1399 LGSSEPYEDFLSADSKVLRGRSSEPECKNGKQGFDPDRSPVSDLTSTGLYDQKQEKST 1458
QY 1019 QALQAVADKVEE--SQVPATQVQRTGSKALEKVEEVEEDSEVLASEKEKDVMPKGPVQ 1077
DB 1459 GFI-----PIKEDFGEPEKKTSDVETMSSQSALALDERKLGIDV--SPQIDVSGSFK 1510
QY 1078 EGAHEHLAOGSETGQ-ATPESLEVPVYT-ADVDAHATCOVITKLQQLMEQAVAPESSETLT 1135
DB 1511 EDTKMSISRGTVSDKSAATPDECAVDYSHMGVASV-----STAVATSSFP 1559
QY 1136 DSTNSTPLADSDTADGQOQDETISQDSKATAAARQSOVTEEAAATQKE----- 1188
DB 1560 EPTTDVSSSLNAEV--GSHSTFVD--DLSLSVSVQTPPTFETEMSPKEKCPKPMST 1615
QY 1189 -----PSTLPNNVPAQOEHEGEBGRDV-----LEPTQOELTAAPVY 1225
DB 1616 SPDFSPKTAKSRTPVQDRHSEQSSMSIERGQSPHESFAMDRSPDPHPLTGA--V 1672
QY 1226 LATVEGQGEVWMLDGEKKEQEVFVHSGPMSOKAADVTYP-----SEWGVAGCQEKES 1282
DB 1673 LHITENGPT-EVDY-----SPCIDSSLSHKIPTEEPS-YTQNDLSSELLSVSQVASPS 1727
QY 1283 T-----EVOGSLSEEGEMTDVEK-EKRE 1305
DB 1728 TSSAHTPSQIASPLQEDTISDVPPRPMSTIYASLASEKVSLEGKLSPKSDISPLTPRE 1787
QY 1306 TKP-----EQVSEEGQETAAPEHEGTYGKV-----LTLTD 1336
DB 1788 SSSLVSPGSDTSAAKETAA-AHQASSSPPIDAATAPYGFRRSMLPTMQHLLNRD 1846
QY 1337 MPSS-----ERKALG-----SLGSPSLPDDAKAGCTIEVYQSLDT--TYTQTA 1379
DB 1847 LTTSSVEKDSGGKTPGDFNYAYOKPEVAAGSPDEEDYD-----YESQEKTRTDVVRYYY 1902
QY 1380 EAVEXIETVVISSEGESPPCVAHLLPAEKSSAT--CGHWTLQHAEDTVPLAPESQARS 1437
DB 1903 EKTERTIKSP--CDSGSVYETT-----EKTIKPBGDGYTEIETIKTTRT-PEEGGYS 1952
QY 1438 IPIVTPAPESTLHPDLOGEISASQERSEEE-----EDKPDAGPD-ADGKESTALE 1487
DB 1953 YEL-----SEKTTTPREVSG-YTEKTERSRRLDDISNGVDDEGDGHTLGDOSYSETT 2007
QY 1488 KVLAEPEILEFSKSNKIVLVNIQAVADQFARTETAPEETHAY-----DSQTOVPACRLD 1542
DB 2008 EKITSFP--ESESYS-----YETSTTSTRSPDTSAYACYETEMKITRTPQASTV 2053
QY 1543 SRE-PNRCTKMKDAMKHPVPQPREDLQVLTV 1574
DB 2054 SYETSRCYTERK--KSP-SEARQDVLDCLV 2081

```

RESULT 3
ANK2_HUMAN
ID ANK2_HUMAN STANDARD; PRT; 3924 AA.
AC Q01484; Q01485;

```

DT 01-APR-1993 (Rel. 25, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-OUN-2002 (Rel. 41, Last annotation update)
DE Ankyrin 2 (Brain ankyrin) (Ankyrin B) (Ankyrin, nonerythroid).
GN ANK2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).
RC TISSUE=Brain stem;
RX MEDLINE=91302466; PubMed=1830053;
RA Otto E., Kunimoto M., McLaughlin T., Bennett V.;
RT "Isolation and characterization of cDNAs encoding human brain
RL J. Cell Biol. 114:241-253 (1991).
RN [2]
RP REVISIONS.
RA Carpenter S.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Brain stem;
RX MEDLINE=94075409; PubMed=8253844;
RA Chan W., Kordeli E., Bennett V.;
RT "440-kD ankyrinB: structure of the major developmentally regulated
RL domain and selective localization in unmyelinated axons.";
RN [4]
RP SEQUENCE OF 463-495 FROM N.A.
RX MEDLINE=92009921; PubMed=1833308;
RA Tse W.T., Menninger J.C., Yang-Feng T.L., Francke U., Sahr K.E.,
RA Lux S.E., Ward D.C., Forget B.G.;
RT "Isolation and chromosomal localization of a novel nonerythroid
RL ankyrin gene.";
RN [5]
RX Genomics 10:858-866 (1991).
CC -!- FUNCTION: Attach integral membrane proteins to cytoskeletal
CC elements. Also bind to cytoskeletal proteins.
CC -!- ALTERNATIVE PRODUCTS: 3 ISOFORMS, 1 (SHOWN HERE), 2 AND 3; ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -!- TISSUE SPECIFICITY: PLASMA MEMBRANE OF NEURONS AS WELL AS GLIAL
CC CELLS THROUGHOUT THE BRAIN.
CC -!- PTM: PHOSPHORYLATED AT MULTIPLE SITES BY DIFFERENT PROTEIN KINASES
CC AND EACH PHOSPHORYLATION EVENT REGULATES THE PROTEIN'S STRUCTURE
CC -!- SIMILARITY: CONTAINS 23 ANK REPEATS.
CC -!- SIMILARITY: CONTAINS 1 DEATH DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X56957; CAA40278.1; -
DR EMBL; X56958; CAA40279.2; -
DR EMBL; Z26634; CAA42644.1; -
DR EMBL; M37123; AAA62828.1; -
DR PIR; S14533; S14533.
DR PIR; A39643; A39643.
DR PIR; B39643; B39643.
DR PIR; S14569; S14569.
DR HSR; P42771; IDC2.
DR Genew; HGNC:493; ANK2.
DR MIM; 106410; -
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000488; Death.
DR InterPro; IPR000906; ZUS.
DR Pfam; PF00023; ank; 24.
DR Pfam; PF00531; death; 1.

```


OY	930	TQTLPEGKDADHDVMYTSVEVPTSEAVTATETSAALTE-----	EYTBAS	973
Db	2685	-YTRKAMEDTQEEFGKSEEEKDSBSHLA--EDRHAVSTEADERSYDKLNDRDTPQPKICGH	2742	
OY	974	GAETITDMVSA--VSQLTDSPP-----	DTTEATPYQVEESG	1007
Db	2743	GCEAMSPSSSARRPVSSQLQSPFTGDDVDQGVITYKSLAQTHKDTSGELBUDSRASRP	2802	
OY	1008	VLDTEEBERQTQALIQ--AAVADKYKEBSQVPATQTVQRTG--SKALEKVEEVED-----	1058	
Db	2803	QADCPSEFSFSSSSSLPCLVBSGKELDEDISATSSIQKTEVTKTETENLPKQCPQSODS	2862	
OY	1059	-----SEVLASEKKKVMKPGPQVEAGAHLAGSETGQATP-----E	1096	
Db	2863	SITQTQDRFSMDVDPVSLAENDELYDPQITSPLENVPSQFSFSSSEKSTQTDPAHNTTSFTH	2922	
OY	1097	SLEVEPV--TADVDHVATC-----OVILQLOOMEQAVAPESSENTLTDSE	1138	
Db	2923	SSEVYSVITISPVEDVNVVASSSGTVLSKESNPEQDIMEQSLSTL--WENQOSV	2978	
OY	1139	TNGSTP-LADSDTADGTQODETIDSQ--DSKATAAVR-----OSQYTEE---AATA	1184	
Db	2979	SSSEPTMTSATTTVVGQIQIKVITITKTDVDSMSIREDDFAFARVVEEOKIFGLMV	3038	
OY	1185	QKEBESTLPNNVPQOEHEGHEPGRDVLPEPQGL--TAAAVPLAKT---EVQGEQEDVM	1239	
Db	3039	DRQSQTTPDPTTPAR-----TPREESTPSEONPLPFGGLFEMTKSGALD--	3085	
OY	1240	LDGEVKEQEEQVFVHSGPNSQKADADVTYDESVW--GVAGCOEKE--STEVOSLSLEGEMET	1297	
Db	3086	MTKRSYADESHFPIQIG--QESREETLSBDVKEGATGADPLPLFSAESLALSSEKETV	3142	
OY	1298	DVEKREKRETKPEQVSEGEQ-----ETAPHEGTYGKPVLTLDMPSE	1341	
Db	3143	D--DEADLLPDSVSEEBEPIPASDAQNSQMSISASTETPKENV---SVGTQDLPTVQ	3196	
OY	1342	RGKALGSLIGG--SPSLPDQKAGIEVQVQSLDTVYQTAABAVEKIEFTVVISENG--ESP	1398	
Db	3197	TGD-IPLISGVKQIISCPDSEPA---VQVLDLDTSTLRS-----YSDNGDDSP	3241	
OY	1399	ECVGAHLLPAEKSSATGCHMTLQHAEDTVPLG-----ESQASBIRPI---IVTPAPE	1447	
Db	3242	DS-----SPEQKQSVI-----EIPTAEMENVPPFESKSK--IPKTIWPTSTPAP	3284	
OY	1448	STLHPDLOGELISAQORSEBEDKPDAGPADGKESTALEKVL-KAPEPILEBSKSKI	1506	
Db	3285	SA--EYESSVSEDFLSSVDEENAD---EAKPKSLPKVPLQREQOQLSDLTSTVQKT	3338	
OY	1507	VL---NVICAVQAPARTETAPETAPETHAYDSQTOVPACRLDSERPNCWTMKMDAMKRPV	1562	
Db	3339	VAPQGDMASTAPPNRKSES--DASLDSKTCIPV-----KTRSYETETESRER--	3387	
OY	1563	POPREDLQVLTLEAMAQPRKCLPRLQJLKA	1592	
Db	3388	---AEELE-LESEEGATRPILTSRLPVS	3413	
RESULT 4				
MAPB_HUMAN		STANDARD;	PRT;	2468 AA.
AC	P46821;			
DT	01-NOV-1995 (Rel. 32, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Microcubule-associated protein 1B (Map 1B) [contains: MAP1 light chain			
DE	LCI]			
GN	MAP1B.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
NN	[1]			

```

RP SEQUENCE FROM N.A.
RC TISSUE=fetal brain;
RX MEDLINE=95104835; PubMed=7806212;
RA Lien L.-, Feener C., Fischbach N., Kunkel L.M.;
RT "Cloning of human microtubule-associated protein 1B and the
RT identification of a related gene on chromosome 15.";
RL Genomics 22:273-280(1994).
CC -I- FUNCTION: THE FUNCTION OF BRAIN MAPS IS ESSENTIALLY UNKNOWN.
CC PHOSPHORYLATED MAP1B MAY PLAY A ROLE IN THE CYTOSKELETAL CHANGES
CC THAT ACCOMPANY NEURITE EXTENSION. POSSIBLY MAP1B BINDS TO AT LEAST
CC TWO TUBULIN SUBUNITS IN THE POLYMER, AND THIS BRIDGING OF SUBUNITS
CC MIGHT BE INVOLVED IN NUCLEATING MICROTUBULE POLYMERIZATION AND IN
CC STABILIZING MICROTUBULES.
CC -I- SUBUNIT: 3 DIFFERENT LIGHT CHAINS, LC1, LC2 AND LC3, CAN ASSOCIATE
CC WITH MAP1A AND MAP1B PROTEINS.
CC -I- DOMAIN: Has a highly basic region with many copies of the sequence
CC KKKE and KKKE/V, repeated but not at fixed intervals, which is
CC responsible for the binding of MAP1B to microtubules.
CC -I- PTM: LC1 is coexpressed with MAP1B. It is a polypeptide generated
CC from MAP1B by proteolytic processing. It is free to associate with
CC both MAP1A and MAP1B. It interacts with the amino-terminal region
CC of MAP1B (By similarity').
CC -I- SIMILARITY: TO MAP1A.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; L06237; AAA18904.1; -.
DR GENE; HGNC:6836; MAP1B.
DR MIM; 157129; -.
DR InterPro; IPR000102; MAP1B_neuraxin.
DR Pfam; PF00414; MAP1B_neuraxin_10.
DR ProSite; PS00230; MAP1B_NEURAXIN; 6.
KW Microtubules; Repeat; Phosphorylation.
FT CHAIN ? 2468 MAP1 LIGHT CHAIN LC1.
FT REPEAT 1878 1894 MAP1B 1.
FT REPEAT 1895 1911 MAP1B 2.
FT REPEAT 1912 1928 MAP1B 3.
FT REPEAT 1929 1945 MAP1B 4.
FT REPEAT 1946 1962 MAP1B 5.
FT REPEAT 1963 1979 MAP1B 6.
FT REPEAT 1997 2013 MAP1B 7.
FT REPEAT 2014 2030 MAP1B 8.
FT REPEAT 2031 2047 MAP1B 9.
FT REPEAT 2048 2064 MAP1B 10.
FT DOMAIN 589 790 LYS-RICH (HIGHLY BASIC, CONTAINS MANY
KKEE AND KKEI/V REPEATS).
SQ SEQUENCE 2468 AA; 270618 MW; 540839CDBD09D461 CRC64;
Query Match 6.1%; Score 492; DB 1; Length 2468;
Best Local Similarity 20.3%; Pred. No. 4.9e-10;
Matches 368; Conservative 266; Mismatches 694; Indels 482; Gaps 75;
QY 45 ADPAKKL-----PCKNGQLSVNGVAEGDVH-----VQENNGSGEEFVDDEVGORE 93
Db 478 ANPAKKIRIVLPFGNSTYNILLEGELKHNDLFKLQPLATOKDLGGOVPPTPVAKTKRKQ 537
QY 94 SEDVEKDORVEEMANSTAVEDITDGOGSETSIEIQIPASENNVNEVMQAPASANDVG 153
Db 538 RADSHRS--LKPAKKPLPSKVKESKEETFEV---TKVNHE--KPKVESKE-- 584
QY 154 FKVKVFVGFKFTVKDKNKSVDYQLLTIVKKDGEGBAASVAGADHOEPS-VETAVGES 212
Db 585 -----KMVKKDKPVKTETKPSVTXKE-----VSKKEPSPVKAELVAE- 622
QY 213 ASKESELKOSTEKQEGTLKQEOSSTEFLPLOAESDOABEAADBECEKOKEEPTSPESP 272
Db 623 -----KQADTVKPKAAKETVKKETEIVKREDDKKEEKEKPKEVAKKEDKPIKKEEP 675

```


RT "Neuraxin, a novel putative structural protein of the rat central
RT nervous system that is immunologically related to microtubule-
RT associated protein 5." ;
RL EMBL J. 8:2879-2888(1989).
RN [4]
RP DEVELOPMENTAL STAGE. AND PHOSPHORYLATION.
RX MEDLINE=97405699; PubMed=9260743;
RA Ma D., Nothias F., Boyne L.J., Fischer I.,
RT "Differential regulation of microtubule-associated protein 1B (MAP1B)
RT in rat CNS and PNS during development." ;
RT J. Neurosci. Res. 49:339-332(1997).
CC -I- FUNCTION: The function of brain MAPs is essentially unknown.
CC Phosphorylated MAP1B may play a role in the cytoskeletal changes
CC that accompany neurite extension. Possibly MAP1B binds to at least
CC two tubulin subunits in the polymer, and this bridging of subunits
CC might be involved in nucleating microtubule polymerization and in
CC stabilizing microtubules.
CC -I- SUBUNIT: 3 different light chains, LC1, LC2 and LC3, can associate
CC with MAP1B and MAP1B proteins.
CC -I- TISSUE SPECIFICITY: Nervous system (spinal cord, brain stem,
CC cerebellum and cerebrum). Not expressed in liver, spleen, kidney,
CC heart or muscle.
CC -I- DEVELOPMENTAL STAGE: In cerebral cortex, spinal cord and sciatic
CC nerve levels are high early in development but decrease during
CC postnatal development and are low in adults. In dorsal root
CC ganglia levels remain high throughout development.
CC -I- INDUCTION: By nerve growth factor.
CC -I- DOMAIN: Has a highly basic region with many copies of the sequence
CC KKEE and KKEI/V. repeated but not at fixed intervals, which is
CC responsible for the binding of MAP1B to microtubules.
CC -I- PTM: LC1 is coexpressed with MAP1B. It is a polypeptide generated
CC from MAP1B by proteolytic processing. It is free to associate with
CC both MAP1A and MAP1B. It interacts with the amino-terminal region
CC of MAP1B (By similarity).
CC -I- PTM: Phosphorylated.
CC -I- SIMILARITY: TO MAP1A.
CC -I- CAUTION: A C-terminal fragment of this protein (residues 1597 to
CC 2459) was originally described as neuraxin in Ref.3.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to license@isb-sib.ch).
CC -----
CC
DR EMBL: U52950; AAB17068.1; -.
DR EMBL: X60370; CAC16162.1; -.
DR EMBL: X16623; CAA34620.1; ALT_SEQ.
DR PIR: S06017; S06017.
DR InterPro: IPR000102; MAP1B neuraxin.
DR Pfam: PF00414; MAP1B neuraxin; 10.
DR PROSITE: PS00230; MAP1B NEURAXIN; 8.
KW Microtubules; Repeat; Phosphorylation.
KW CHAIN ? 2459
FT REPEAT 1869 1885 MAP1B 1.
FT REPEAT 1886 1902 MAP1B 2.
FT REPEAT 1903 1919 MAP1B 3.
FT REPEAT 1920 1936 MAP1B 4.
FT REPEAT 1937 1953 MAP1B 5.
FT REPEAT 1954 1970 MAP1B 6.
FT REPEAT 1988 2004 MAP1B 7.
FT REPEAT 2005 2021 MAP1B 8.
FT REPEAT 2022 2038 MAP1B 9.
FT REPEAT 2039 2055 MAP1B 10.
FT DOMAIN 559 1035 GLU-RICH.
FT DOMAIN 588 786 KEE-RICH (HIGHLY BASIC, CONTAINS MANY
FT LYS-AND KEEI/V REPEATS).
FT DOMAIN 2224 2312 LYS-RICH.
FT CONFLICT 127 127 M -> V (IN REF. 1).
FT CONFLICT 140 140 T -> S (IN REF. 1).
FT CONFLICT 2112 2112 R -> K (IN REF. 3).

FT CONFLICT 2169 2169 L -> I (IN REF. 3).
SQ SEQUENCE 2459 AA; 269497 MW; 2E3F6872DEDB8BA2 CRC64;
Query Match 5.8%; Score 470; DB 1; Length 2459;
Best Local Similarity 20.2%; Pred. No. 2.8e-09;
Matches 328; Conservative 234; Mismatches 602; Indels 458; Gaps 62;
QY 45 ADPAATL-----PQKQGLSSVNGVAEGDVH-----VQENQEGQEEVVDVQORE 93
DB 477 ANPAAKRIIRVLPFNSQSYNLEGLLEKILDLFLKQPLATQKDLTGVSPTPPVQVNLKQ 536
QY 94 SEDRE--KDEVEEAMNSTAVEDITKQGESETIIEQIPASENNVEWQAPASQAND 151
DB 537 RADRSRLKAPTKPLSSKS-----VRKESKEAE--ATASQVEKTPKESKE-- 583
QY 152 VGFKKVPKFPVFKTKKDKNEKSDTVQLTVKKDEGEBAASVAGADHQPVSVEVAVGE 211
DB 584 -----KVIYKKDKPKGVESKPSVT-----EKEVPSKEQSP-VKAEVAE 621
QY 212 SAKSESELKQSTKQEGTLKQEGSTELPLQAESDQAAEEAKDEGEKQKEPTKSPES 271
DB 622 KAATESKPKVTQDK--VVKKE-----IKTKPEKKEKPKKEVAKKEDKTLKKDEK 671
QY 272 PSSPVSSTTSFPKKFTFHGAWGRKKTSPKSKEDDLETAKEKKEQAEKVEEEK--E 329
DB 672 PK--KEAAKEIKK-----EIKKEKEKELK-KEVKEITPLKDAKKEVAKDE 714
QY 330 KTEPASEBOEPAEDTDQALSLADYKVELPLEDQVGDLEASSEKCAPLATEVDEKMEA 389
DB 715 KKEVKEKEKEPKKEI--KKISKDIKK--STPLSD-----TKPKALKPKVAKKEEPT 762
QY 390 HQEVAAEHVAVTEKTEEGOGGGAEGVAVVEGTGSELTPREKLAEPQVQEAPEAEEL 449
DB 763 KKEPIAAKLDKDKGVKVKYKKEGTTETAATAVAGTAALAAAGAA-----AGPAEEL 815
QY 450 MKSREMCVSGGDHTQ-LTDLSPKEKTLPKHPEGIVSEWMLSSQERIKVQSGPLKLFSS 508
DB 816 EAERSLMSPEDLTQDFELAAEELIDVAK--DIKQLELEDEBKILK----- 860
QY 509 SGLKKLSGKKQKGRKGGGDEPGEYOHINTESPESADEQKSSASSAPSEPEETTCLEK 568
DB 861 -----ETEPGEAVYIOKE-----TEVSK 878
QY 569 GPLEAPQGEABEGTTSQGEKKREGITPMASFMMVTPKAVRPSBDEKEELKVKSA 628
DB 879 GSABSPDDEGI--TTTEGESECE-----QTP-----EELEPV-- 907
QY 629 TLSTSTVSEMDQEVKTVGEQKPEEPKRRVDSVSNEALICVSSKKRARKASSDDE 688
DB 908 -----EKQGVDDIEKTEDEBAGCFEBSSEADYIEKA----- 938
QY 689 GGPRTLAGDSHRAERASKDKAGTDVAVPASTOEDQAOAGSSPEPAGSPSEGEVSTWES 748
DB 939 -----ETBEAREPEEDG-----EDNVGSASAKH--SPTDEBETAKAAEA 974
QY 749 FKALVTPPKKSKSKLEEAADBSVQLST-ELPEPESWSWISIKFTIGRKKRADQGE 807
DB 975 DVH--IKERRESVYASDDRAEDMDALEKGEAEOSEEBG-----EEBEDVAE 1020
QY 808 QATVEDSGPVEINEDDPVAVVPLSEYNAAVEREKGAEQNTLPQLLGAVVYSELSKT 867
DB 1021 DAREEDHEPDTEARD-YVMAVVDKAAAGATVEDQYDLGT--PAKPPQOSPEREAS 1076
QY 868 LVHTVSVAVIDGTRAVTSVEER--SPSWISAS--VTEPLEHTAGAMPVEEVEKDI 922
DB 1077 SIHDETLPGGSESEATADENREDDQPEFPATSGYTGSTIEISEPPTMDMSPTPRVM 1136
QY 923 AEET-----PYLT--QTLPECK--DAHDMVTSVD 949
DB 1137 TDETNNETESPSPQFVNIKYESSLYSQEYSKPVVASFNGISLDSKTDATDGR-----D 1191
QY 950 FTSEAVTATENS-----EALRTEVTVASGAETTDVMS-----AVSQLT 989


```

Db 1192 YNASASTISPPSMEDKFKSALROAYRPEETDVKTAGELDIKDVSDERLSPAKSPSL 1251
QY 990 DSPDTTEATPVORVSGVLTDEERQTOAILQAVADKVKESQVPATVTVORTGSKAL 1049
Db 1252 PSPSPTEKTPLGERSVNFSLTNE-----IKASAEGEATAVSPGVTOAV 1297
QY 1050 -----EKVEEVEDSVLASEKEKVMKPGVQVQEAHAQSGEQATPESLEV 1100
Db 1298 VEHCASPEEKTLEVVSPQSOTVSGAGHTPYQSPDEKSHLPTVETENAAQVPSVFEF 1357
QY 1101 PEVTADVDHVCATC-----QVTKLQQLMEQAVAPESSETLTDTSETNGSPPLADSDTADCTQ 1156
Db 1358 TEAKDENERSSIPMDPEVDSPIEKVLSPLRSPPLIGSEAYEDFLSADKALGRS 1417
QY 1157 DETIDQDSKATAAVQSQVTEBEAATAQKEPSTLPNNVPAQEEHGEPRGVLEPTQ 1216
Db 1418 ESPFEKNGKQGFSDKESPVSLDLSLYQDKBEKAGFTPIKEDFSPEKKASDAEIMSS 1477
QY 1217 ELTAAAVPVLAKEVGOEG---EVDMLDGEKVKEEQEVFVHSGPNSOKA-----ADVT 1266
Db 1478 QSALA-----LDERKJGDGSPQTQDVSVQFGSKFEDTKMSISEGTVSDKATPVDEGAEDT 1533
QY 1267 YDSEVMVAGACQ-----EKESTEVEQSLSLEGE----- 1294
Db 1534 Y-SHMEGVASVSTASVATSFPEPTDDVSPSLHAEVGSPHSTEVDDLSVSVVQTPTF 1592
QY 1295 METDVKKEKE-----TKPEQVSEGEQETAAPHEGYGKPVLTLDMPSSERKALG 1347
Db 1593 QETEMSPSKECPKPMGSPDFSPKTAKSRTPVQDHRSPQSSMSIEFGQESPEHSLAMD 1652
QY 1348 SLGGSPSLPDQKAGCT-----EVQVQSLDTTQVTOAEAV-----EKVIET 1388
Db 1653 FSRQSPDHPTVG-AGMLHITENGTEVDYSPSDIQDSSLKHPPTPEPSYTDNDLSEL 1711
QY 1389 VITSETGESPECVCAHLLPAEKSSATCGHWTLOHAEDTVPLGPESQAESPIIIVTPAPES 1448
Db 1712 ISVSQVEASPTSSAH-TPSQIAS-----PLQEDTLSD-----VPPRDM 1750
QY 1449 TLHPDLOGE-ISASORER-----SEEDKPDAGP-----DADGKESTAIEKVLKA 1492
Db 1751 SLVASLASEKQVSLGEGKLSPKSDISPLTPRESSPTVSPGFSSTSGAKESATAAYQTSSS 1810
QY 1493 EP 1494
Db 1811 PP 1812

RESULT 6
MAPA_RAT
AC P34926;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Microtubule-associated protein 1A (MAP 1A) [Contains: MAP1 light chain
  LC2].
GN MAPA.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=92355629; PubMed=1379599;
RA Langkopf A., Hammarback J.A., Mueller R., Vallee R.B., Garner C.C.;
RT "Microtubule-associated proteins 1A and LC2. Two proteins encoded in
  one messenger RNA."
RL J. Biol. Chem. 267:16561-16566(1992).
CC -!- FUNCTION: Structural protein involved in the filamentous cross-
  bridging between microtubules and other skeletal elements.
CC -!- SUBUNIT: 3 different light chains, LC1, LC2 and LC3, can associate
  with MAP1A and MAP1B proteins.

```

```

CC -!- TISSUE SPECIFICITY: BRAIN, HEART AND MUSCLE.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED LATE DURING NEURONAL DEVELOPMENT
  APPEARING WHEN AXONS AND DENDRITES BEGIN TO SOLIDIFY AND STABILIZE
  THEIR MORPHOLOGY.
CC -!- DOMAIN: THE BASIC REGION CONTAINING THE REPEATS MAY BE RESPONSIBLE
  FOR THE BINDING OF MAP1A TO MICROTUBULES.
CC -!- PTM: VARIOUS SERINE RESIDUES MAY BE PHOSPHORYLATED BY CAMP KINASE.
CC -!- PTM: LC2 IS COEXPRESSED WITH MAP1A. IT IS A POLYPEPTIDE GENERATED
  FROM MAP1A BY PROTEOLYTIC PROCESSING. IT IS FREE TO ASSOCIATE WITH
  BOTH MAP1A AND MAP1B.
CC -!- SIMILARITY: TO MAP1B.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M83196; AAB48069.1; -.
CC PIR; A43359; A43359.
CC CHAIN 22465 2774 MAP1 LIGHT CHAIN LC2.
CC FT DOMAIN 309 496 LYS-RICH (BASIC).
CC FT DOMAIN 336 541 11 X 3 AA REPEATS OF K-K-[DE].
CC FT REPEAT 336 338 1.
CC FT REPEAT 415 417 2.
CC FT REPEAT 420 422 3.
CC FT REPEAT 424 426 4.
CC FT REPEAT 427 429 5.
CC FT REPEAT 431 433 6.
CC FT REPEAT 436 438 7.
CC FT REPEAT 440 442 8.
CC FT REPEAT 444 446 9.
CC FT REPEAT 449 451 10.
CC FT REPEAT 539 541 11.
CC SQ SEQUENCE 2774 AA; 299526 MW; 3DEF74427BA9D7D7 CRC64;

Query Match 5.4%; Score 433.5; DB 1; Length 2774;
Best Local Similarity 22.0%; Pred. No. 5,5e-08;
Matches 388; Conservative 231; Mismatches 625; Indels 523; Gaps 89;

QY 103 VEEMAANSTAVEDIT-KDQGE-----ETSEITQIIPASNNVEMVQ---PAESQA 149
Db 216 MQWAGNSKAKTGVLANGKEAEISVPYLTSITALLVWLPA--NPTEKIVRVLPFGNAPQ 273
QY 150 NDV--GFKV--FKFVGPKFTVKDKVKSQSDTVLLTVKKDEGEAEASVAGDHOEPSV 205
Db 274 NKILEGLEKRLHLDLFLRYPVATQKDLAAGAVPANLKPSKIKHRADSKESLKA-----AP 327
QY 206 ETAVGESASKESELKQSTKQSGTLKQEOSSTEIPLQAESDQAAEEAKDEGEKEKEP 265
Db 328 KTAVSKLAKREVLLEGAKESLAKELAKT-----EKKAKEP 366
QY 266 T-KSPESPSP-----VNSETTSSPKFPTHGWAGWRKTKTSFKKSKEDDLETAEEKEQE 319
Db 367 SEKPPKSPKSERVGSSEALKAERLIKQAG-----KXHLKEKISKLEKKDKE 419
QY 320 AKVDDEEKE-KTEPASEQEPADTDQARLSADYKVELPLEDQVGDLEASSEEKCAPL 378
Db 420 KKEIKKERKELKEGRKEKKDAKKDKR-----KDTKPEVKLSKPLDKPF 467
QY 379 ATEVPDE-----KMEAHQEVVAEVHVSTVEKTEEEOGGGGAEAGGVVVGTSGL 428
Db 468 TPEVRKTLVYKAPGRVKVDKGRAARGEXELSEPRTPPAQKG---AAPPAAVSGHREL- 523
QY 429 PPEKLAEPQEVQEAPEABELMKSRMCMVSGGDHTLTDLSPEEKTLP----- 476
Db 524 ---ALSPEDLTQD---FEELKREER-----GLLAQRDTGLGEKPLPADATEQHPASA 572
QY 477 ---KHPEGTIVSEVEMLSQRIKVGQSPILKLFSSSGLKKLSGKKQKRGKGGGDEEPE 533

```

```

Db 573 IGVTPSGVLEGEHVEREKEY-VPDSDPKD-----GSTNRGPDGSAVEYKEKE 620
Qy 534 YOHITHTEPSADECQKSSASSPEPE---ETTCLKQPLFA-----PDGEAEGETTS 585
Db 621 ---TWBRKQREALGENTTAARESEAEVEDVIEKALEEMETHPSDEGEETAE 676
Qy 586 DGEKK--REGITPMASFVKMTYPRKVR-----RSESEKEELEKVSATLSSTD 635
Db 677 SFYQKHTGEL-----KASPKREALGGRDLGFQKAPKETAFLSLSTAGATE 728
Qy 636 TVSEKODEV-----KTVGEQKPEEPKRVDTSVSEWALICVSSSKRARKSSSD 687
Db 729 HVSYIQDETIPEYSETEQIISDEIHDEPDER-----PAPRPFSTYTD 772
Qy 688 EGGPRTLLGDSHRAEASKDEAGTDAVPAS---TQEDDQAGSSSPPEAGSP--SEGEV 743
Db 773 LSGPEGPG-----PREAS---QAADSAVPASSSKTYGAPETELTYPPMVAAPLAEETHV 824
Qy 744 STWESFKRLVTPRKSKSKLEKADSSVEQLSTELPERSRWVS-----IKKFIPIGR 798
Db 825 SSATS-----ITECDKLSSPATSVAEDOSVASLTA---PQTEBTGKSSLLDVTGSIPISSR 877
Qy 799 KKBAGKQEQALVEDSGPEINEDDPNPAVVPPLSERYNAVEREKPEAQNTLIPOLLGA- 857
Db 878 TEATQGLDVPESAGTISPTSSLEBDKGFKS--PCBEDPSVTGES--EKGETVGRGLSGEK 934
Qy 858 ---VVYSEELSKTLV-----HTVSVAVIDGTRAVTSVEER--SPSMISAVTE 900
Db 935 AVGREKXYVNSEKLSQGVAAVFGAPGHTLP--PGEPALGEVERCLSPDSTVYKMS 990
Qy 901 P-----LEHTAGAMPVEVETK----- 919
Db 991 PPPSGPPSAHATPHQSPVEDKSEPRDFQEDSGWGTGKSHPGVSKEDSEQTVKPGEEGT 1050
Qy 920 ---DIAETPVLTOTLPE-----GDADHDWMTSEV 948
Db 1051 SREGGPPTRSQADMPVSLGGQTCITQLLPEDDKAIVETGAGNLAGTLPGEV 1110
Qy 949 DTSBAVTATETSEALRTEETVTEASCAETTMVSAVSQLTDSPTT--EATPVQEVESG 1007
Db 1111 RSTTEATPOQDEVLR---FTDQSLSPEDASLSVLSV--SPDTTKQEAIP-----RS 1160
Qy 1008 VLDTEEE-----KOTQALIQAVADKVESQVPAQTVQRTSKSL--EKVEE 1054
Db 1161 PCSLEEQQPHKMLMPVSPEDTQSL-----SFSEBS--PSKRTSLDISKSQSPESLGT 1212
Qy 1055 VEDESEVLASEKQVMPKPGVQEAAG--HLAQGS--ETGAT--PESLEVP----- 1101
Db 1213 IQFGELINIGKEE---NGPVWKAEDDSCHLAPVISPEPHRATVSPSTDETPAGTLPGG 1266
Qy 1102 ---EYADVDHVATCQVILQQLMEQAVAPESSETLTDSETNGSTPL 1145
Db 1267 SFSHSLSVDRKSHPEITGPGHEMT-----SDSLSLTKSPESLSIPA 1309
Qy 1146 AD-----SDPADTQOQDETIDSQSKATAVAVROSQVTEEAATACK-----EPSTLP 1193
Db 1310 MEDLAVEMEGKAPGKEKEBELKSETRQ-----QKQIILPEKVAVVEDDILIHOKGALDE 1364
Qy 1194 NNVPV--QEHHEEPGRDVL--PTQOELTAAAVPVLAKTEVQEGEVDMLDQKVEBOE 1250
Db 1365 ENKPGQDQKTEPQKGRDIDEXDTAELDKGPP-----KEXLD-----REDG 1409
Qy 1251 VEVHSGPNQSK--AADTVYSEWGVAGCOEKESTEVQSLSEGEW--ETDVEKEKETRP 1308
Db 1410 ---QRAQPAEKQKASQRTDL-----QOTQATEPRDAQGERDSEKDKSLFLDRTP 1461
Qy 1309 EGVSEGEQETAAPHEGTGYKPVLTIDMP-----SESEKAL-----GSIG 1350
Db 1462 BEKDRILVQEDBAPEN--SIPEPTQDRAPDRKGTDEKQKEBESPEKQVLEQKQWALG 1519
Qy 1351 GSPSLPDQKACIGVQVQSL--DTTVTQTAKEVKVI---ETVVISENGESPECV--- 1401
Db 1520 KEGETLDOE--ARTAEQKDETLKEDKTQOGKSSFVEDKTKTTSKETVLDOXSABEADSVEQ 1578

```

```

Qy 1402 -GAIL-----LPAKSSATGS-----HW-----TLQHAEDTVPLGR---SQAESIP 1440
Db 1579 DGAALFKTRALGLEESPESKAREQKRYKWEQDVQGMRETSPTGEPVGGKEPVPA 1638
Qy 1441 IVTPAPESTLH-----PDLQGISASORERS-----EEDKPDAGPD--ADGKESTAE 1487
Db 1639 WEGSPQEVAVYWRDRITLLOQDAYWRELSCDRKVWPPEHLDQGARPRYCEBERESTFLD 1698
Qy 1488 KVLAAPEPILEESKSNKIVLVNIQTAVDQFARTETAPETHAY--DSGTQVPACRLDSREP 1546
Db 1699 EG-PDEQETPLQ-----HTPRSPWTSDFQDFQEPPLPQKGLAVER- 1737
Qy 1547 NRCWTMKMDACKHPVPQPRREDLOYLT 1573
Db 1738 ---WL-----AESPVGLPPEEDKLT 1755

RESULT 7
MAPA HUMAN
ID MAPA HUMAN STANDARD; PRT; 2805 AA.
AC P78559; Q15882; Q12973; Q9UT74; O95643;
AT 15-JUL-1998 (Rel. 36, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Microtubule-associated protein 1A (MAP1A) (Proliferation-related protein p80) [Contains: MAP1 light chain LC2].
GN MAP1A.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97001161; PubMed=8812494;
RA Fink J.K., Jones S.M., Esposito C., Wilkowi J.;
RT "Human microtubule-associated protein 1a (MAP1A) gene: genomic organization, cDNA sequence, and developmental- and tissue-specific expression.";
RL Genomics 35:577-585(1996).
RN [2]
RP SEQUENCE OF 1-1825 FROM N.A.
RX TISSUE=Ovarian carcinoma;
RA Chen Z.C., Fadiel A., Natfolin F.;
RT "Identification of a novel protein (p80) in ovarian carcinoma cells.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 78-1687 FROM N.A.
RX TISSUE=Brain;
RA Ohnani K., Rutherford T., Sakamoto H., Natfolin F.;
RT "Microtubule associated protein 1A (MAP1A) in human brain - DNA sequence and physiological role.";
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 134-419 FROM N.A.
RX TISSUE=Fetal muscle;
RA Chamilikhal N., Paeturaud P., Richard I., Auffray C., Beckmann J.S.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 1607-1883 FROM N.A.
RX TISSUE=Brain;
RA Fukuyama R., Rapoport S.I.;
RL "Brain-specific expression of human microtubule-associated protein 1A (MAP1A) gene and its assignment to human chromosome 15.";
RL J. Neurosci. Res. 40:820-825(1995).
CC -I- FUNCTION: Structural protein involved in the filamentous cross-bridging between microtubules and other skeletal elements.
CC -I- SUBUNIT: 3 different light chains, LC1, LC2 and LC3, can associate with MAP1A and MAP1B proteins.
CC -I- TISSUE SPECIFICITY: BRAIN.
CC -I- DOMAIN: THE BASIC REGION CONTAINING THE REPEATS MAY BE RESPONSIBLE

```

CC FOR THE BINDING OF MAP1A TO MICROTUBULES.
 CC -!- PTM: VARIOUS SERINE RESIDUES MAY BE PHOSPHORYLATED BY CAMP KINASE.
 CC -!- PTM: LC2 IS COEXRESSED WITH MAP1A. IT IS A POLYPEPTIDE GENERATED
 CC FROM MAP1A BY PROTEOLYTIC PROCESSING. IT IS FREE TO ASSOCIATE WITH
 CC BOTH MAP1A AND MAP1B.
 CC -!- SIMILARITY: TO MAP1B.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U38291; AAB41132.1; -;
 CC EMBL; U38292; AAB41133.1; -;
 CC DR EMBL; AF200415; AAF08305.2; -;
 CC DR EMBL; U80458; AAD00355.1; -;
 CC DR EMBL; Z47038; CAA87104.1; -;
 CC DR EMBL; U14577; AAB81362.1; -;
 CC Genew; HGNC:6835; MAP1A.
 CC MIM; 600178; -;
 CC KW Microtubules; Repeat; Phosphorylation.
 CC FT CHAIN 22490 2805 MAP1 LIGHT CHAIN LC2.
 CC FT DOMAIN 309 496 LYS-RICH (BASIC).
 CC FT REPEAT 415 541 9 X 3 AA REPEATS OF K-K-[DE].
 CC FT REPEAT 420 422 1.
 CC FT REPEAT 427 429 2.
 CC FT REPEAT 431 433 3.
 CC FT REPEAT 436 438 5.
 CC FT REPEAT 440 442 6.
 CC FT REPEAT 444 446 7.
 CC FT REPEAT 449 451 8.
 CC FT REPEAT 539 541 9.
 CC FT CONFLICT 72 72 F -> L (IN REF. 1).
 CC FT CONFLICT 134 135 VV -> IP (IN REF. 4).
 CC FT CONFLICT 249 249 A -> G (IN REF. 4).
 CC FT CONFLICT 263 263 V -> A (IN REF. 4).
 CC FT CONFLICT 296 296 Q -> H (IN REF. 3).
 CC FT CONFLICT 311 311 S -> G (IN REF. 4).
 CC FT CONFLICT 324 324 K -> Q (IN REF. 3).
 CC FT CONFLICT 335 336 AK -> ST (IN REF. 1).
 CC FT CONFLICT 353 353 A -> S (IN REF. 1).
 CC FT CONFLICT 357 357 A -> S (IN REF. 1).
 CC FT CONFLICT 364 364 K -> Q (IN REF. 1).
 CC FT CONFLICT 414 419 EKKDKE -> KKKRNS (IN REF. 4).
 CC FT CONFLICT 424 424 K -> P (IN REF. 1).
 CC FT CONFLICT 426 426 E -> D (IN REF. 3).
 CC FT CONFLICT 431 431 K -> Q (IN REF. 3).
 CC FT CONFLICT 439 439 E -> D (IN REF. 3).
 CC FT CONFLICT 444 444 K -> R (IN REF. 3).
 CC FT CONFLICT 452 453 TK -> SS (IN REF. 3).
 CC FT CONFLICT 457 457 K -> R (IN REF. 3).
 CC FT CONFLICT 682 682 Q -> P (IN REF. 1).
 CC FT CONFLICT 1025 1025 Q -> K (IN REF. 1).
 CC FT CONFLICT 1303 1313 KVLPGAITSPD -> EVLTWGDHQAALN (IN REF. 3).
 CC FT CONFLICT 1335 1341 MISSING (IN REF. 3).
 CC FT CONFLICT 1368 1368 Q -> T (IN REF. 3).
 CC FT CONFLICT 1470 1470 A -> T (IN REF. 3).
 CC FT CONFLICT 1650 1650 W -> C (IN REF. 1).
 CC FT CONFLICT 1690 1690 A -> S (IN REF. 1).
 CC FT CONFLICT 1714 1714 G -> V (IN REF. 1).
 CC FT CONFLICT 1869 1869 E -> A (IN REF. 5).
 CC FT CONFLICT 1879 1883 GTPPEY -> AHSRV (IN REF. 5).
 CC SQ SEQUENCE 2805 AA; 306392 MW; F3ED0A3165993B2E CRC64;
 CC
 CC Query Match 5.2%; Score 421; DB 1; Length 2805;
 CC Best Local Similarity 20.8%; Pred. No. 1.5e-07;
 CC Matches 349; Conservative 237; Mismatches 671; Indels 420; Gaps 72;
 CC
 CC QY 103 VEEMAANSTAVEDIT-KDGOE-----ETSEIIIEIQIPASENNVEEMVQ---PAESQA 149

Db 216 MQWAGNSIAKGTGIVLPNGKEAIBISVPYLTSTALVWVLP--NPTEKIVRVLPFGNAPO 273
 QY 150 NDV--GFKKV--FKVGFKFTVKOKNEKSDTVQLLTVKKDEGEAEASVAGGDHOEPSV 205
 Db 274 NKILEGLEKRLHLDPLRYPVATQKQLASGAVPTNLKPSIKQRADSKESLKA-----TT 327
 QY 206 ETAVGESAKSELKOSTKQGTGLKQBSSTEIPLQASDAQAEAEKDEGEKEKEP 265
 Db 328 KTAVSLAKREVEVGAKEARSELAKELAKT-----EKKAKESSE----- 368
 QY 266 TKSPESSPSP--VNSSTTSSPKFFTHGWAGRKKTSPFKSK--EDDLATAKREKEOA 320
 Db 369 -KPEKPAKPERVKTESSEALK-----AEKRLIKOKVGKHLKAKISKULEKKDKK 420
 QY 321 EKVDEEKE--KTEPASEQEPAEDTDQARLSADYEVKLEPLEDQVGDLEASSEKCAPL 378
 Db 421 KEIKKERKELKXDEGRKEEKDA-----KKEEKRDTPKELKTKISKPLKPF 467
 QY 379 ATEVFDEKMEAHQEVVAEVHVSTVTEEBEQGGGAEAGVVVVGTEGSLPPEKLAEPQE 438
 Db 468 TPEV-----RKTLYKAKVGRVKIDRSRAIRGE-----KELSEPO 504
 QY 439 VPQE--AEPAEELMKSRMCVSG-GDHTQ-LTDLSPEEKL--PKHPEGIVSEVEMSSQ 492
 Db 505 PPAQKGTVPPTISGHRELVLSPDLTQDFEEMKREERALLAEQDRTGLGDKPFLDTA 564
 QY 493 ER-----IKVQGGSP-----LKKLFSSSGLKKLQKQKRGKGGGDEPEGEYQH 536
 Db 565 EEGPSTAIQGTTPSPVGLQSEHVMKELVPEVPEEQSGKDRGLDSCAETEE--EKDT 622
 QY 537 IHTESPESADEQKGESSASSPEPE-ETTCLEKGLPEA-----PDGRAEBEGTSGEKK 590
 Db 623 WBEKKQREARLPDRTEAREESEPEVKEDVIEKAELEEMEVEHPVSEDEEDATKAEGFYQ 682
 QY 591 REGITPWASFKKWVTPKKVRPSESDEKEEL-----EKVKSATLSSTDST-----V 637
 Db 683 -----KMQEPLKVTPKSREAFGRELQLOGKAPEKETSLSLFLSLTTPAGATEHV 732
 QY 638 SEMQDEV-----KTVGEQKPEPKRRVDTSVSWEALICVGSKKRARKASDDDEG 689
 Db 733 SYIQDETIPGYSETETISDEIHDEPEER-----PAPRPHSTSYDLP 776
 QY 690 GPRTLGGDSHRAEASKKEAGTAVPASTQDQAGSSSEP-----AGSPSEG 740
 Db 777 GPEGAG-----PFEASQPADS--AVPATS-----GKXAGTPELTPTNIVAAPLAE 823
 QY 741 EGVSTWESFKRLVTPKKSLSKLEAKEDSSVEQLSTEIEPSREESWVS-----IKKFP 795
 Db 824 EHVSSATS-----ITECDLSSFATSVAEDQSASLTA---PQTEETGKSLLLDTVTSP 876
 QY 796 GRKKRADGKQEQATVEDSGPVEINEDDPNVPAVPLSEYNAV-EREKMEAQNTLPOL 854
 Db 877 SSRTEATQGLDVPSAGTISPTSSLEEDGFKS-PCEDFSVTGSEKR----- 924
 QY 855 LGAVVYSEELSKTLVHTVSAVVDGTRAVTSVEERSPS-----WISASVTEPLEHTAG 909
 Db 925 -----GEIIGKGL-----SGRAVEEBEETANVEMSEKLSQYCTPVFSPAGHA 969
 QY 910 MPVPE-----EVTEKDIIAETPV-LTQTLPEKDA--HDDMTSEVDFTSEAVTATET-- 960
 Db 970 LHPGEPALGEAEERCLSPDDSTVKMASPPSPSATHPTFFHQSPVEEKSEFQDFEADS 1029
 QY 961 -SEALRTEBVTASGAETTDMVASVSLTD-----SPDTTEATPVQEVESG--- 1007
 Db 1030 WGDTRKTPGVGKEDAAEETVKFPGPEGTLKEEKKVPPRSPQAQAPVNDIGLCTCTIQ 1089
 QY 1008 -----VLDEEERQTOAILQAVA-----DKVKEESQVPATQVORTGSKALEKVE 1053
 Db 1090 LLPAQDKALVFEIMEAGEPTGILGAEALPGGLRTLPOEPGPKQKDEVLRYPDRSL-- 1146
 QY 1054 EVEEESVLA-----SEKEDVMKPGVQAGAHHLAQGSETQQTATPESLEVPEVTD 1106

```

Db 1147 --PEBAESLSVSVSPSPDTANOEPPTKSPC--GLTEQYIHKDQWREVPSPDTQSLSSSE 1202
Qy 1107 VDHVATQGVYIKLQQLMEQAVAPESSETLTDSNT----- 1140
Db 1203 SPSKETSILVSSKQL-----SPESLQTLQFGEINLCKEEMGHLMQMAEDTSHHTAPMSVPE 1257
Qy 1141 ---GSTPLADBDTADGTQODETISQDSKATPAVQSV-----TEENA 1182
Db 1258 PHAATASPTDGTTRYSAQTDITDSDLRKSPASSSHSTSPGNGKYLPGAITSPEHITL 1317
Qy 1183 TQO-----KEEPTLP-----NNP-----AOEHEGEPGRDYLEPTQOELTAA 1221
Db 1318 TDDSSFSKSPESLPGFALMEDIAIKMEDKVPGLKDRTSSEQKKEPEPDEVLOQKXLTLEHK 1377
Qy 1222 AVPLAKTEVGEQGEVDWLDEKVKKEQEVFVHSGPNSQKADVTYDSEWGVAGCQEKX 1281
Db 1378 EYVEPMDTAIYQKDEALHYKNEAVKQODKALBQK-----RDLBQKDTALEQKDK 1427
Qy 1282 STEVQSLSEEGEMETDVKEKRETPQOVSEGEETAPEHEGTYGKPVLTLPMPSE 1341
Db 1428 ALEPKDKDE---EKDKALEQKDKIPEBKDKALEQKDTALEK---DKALEPKDKDLBQ 1480
Qy 1342 RKGALGSLGSPSLPDPDAGCIEVQVOSLDPTVTQ--TAAVEKYIEFVVISSETGESPE 1399
Db 1481 KRVVLEQKKEIE--EKDKA--LDQKRVSEVHKAPEDTYAEMKDRLE-----QTDKAP 1531
Qy 1400 CVGAHLLEPAEKSSAT-----GGMHTLOHAEDEVPLGPEQSAESIPIVTPAPESTHA 1451
Db 1532 --QKIQAGQKQKVKSEKQKQALQKQKQALQKQKQALQKQKQALQKQKQALQKQKQALQKQKQ 1586
Qy 1452 PLOGEISASQRESEEREKPPADGADGKESATIKV-----LKAPELLEIESKS 1503
Db 1587 ED-----KTRKPKMLEEKSPEVKAMEKLEALLEKTKALGLEESLVOEGRAEQUEEKY 1640
Qy 1504 NRTIVLVICTAADOAFARTAPETAETHAVDSOTQVPACTLSRE--PNCCTKTKMDKAKK 1559
Db 1641 WR-----GQDVVOEWQETSPTREEPAGEOKELAPAMEDTSPEDNRYWNGREVDAL 1692

```

```

RESULT 8
PGCV_CHICK STANDARD; PRT; 3562 AA.

```

```

AC Q90953; Q90945; (Rel. 35, Created)
BT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Version core protein precursor (Large fibroblast proteoglycan)
DE (Chondroitin sulfate proteoglycan core protein 2) (PG-M).
GN CP952.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauromorpha; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OC NCBI_TaxID=9031;
RN [1]
RX STRAIN=White Leghorn; TISSUE=Limb bud;
RX MEDLINE=93300846; PubMed=8314802;
RA Shitamura T., Nishida Y., Ito K., Kimata K.;
RT "cDNA cloning of PG-M, a large chondroitin sulfate proteoglycan
RT expressed during chondrogenesis in chick limb buds. Alternative
RT spliced multiforms of PG-M and their relationships to versican."
RL J. Biol. Chem. 268:14461-14469(1993).
CC -I- FUNCTION: May play a role in intercellular signaling and in
CC connecting cells with the extracellular matrix. May take part in
CC the regulation of cell motility, growth and differentiation. Binds
CC hyaluronic acid.
CC -I- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
CC -I- ALTERNATIVE PRODUCTS: At least 2 isoforms; V0 (shown here) and
CC VI are produced by alternative splicing.
CC -I- TISSUE SPECIFICITY: Prechondrogenic condensation area of
CC developing limb buds.
CC -I- DEVELOPMENTAL STAGE: Disappears after the cartilage development

```

```

CC (By similarity)
CC -I- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
CC -I- SIMILARITY: CONTAINS 2 LINK DOMAINS.
CC -I- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -I- SIMILARITY: CONTAINS 1 SUSHI (SCR) DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC -----
CC EMBL; X60226; CAA42787.1; -.
CC EMBL; D13542; BAA02742.1; -.
CC HSSP; P00740; 1EDM.
CC InterPro; IPR000152; Asx hydroxyl.
CC InterPro; IPR000561; EGF-like.
CC InterPro; IPR000742; EGF_2.
CC InterPro; IPR001881; EGF_Ca.
CC InterPro; IPR003599; Ig.
CC InterPro; IPR003066; Ig_MHC.
CC InterPro; IPR001304; Lectin_C.
CC InterPro; IPR000538; Link.
CC InterPro; IPR000436; Sushi_SCR_CCP.
CC Pfam; PF00008; EGF; 2.
CC Pfam; PF00047; Ig; 1.
CC Pfam; PF00059; lectin_C; 1.
CC Pfam; PF00084; sushi_1.
CC Pfam; PF00193; Xlink; 2.
CC ProDom; PD000918; Link; 2.
CC SMART; SM00032; CCP; 1.
CC SMART; SM00034; CLECT; 1.
CC SMART; SM00179; EGF_CA; 1.
CC SMART; SM00001; EGF_Like; 1.
CC SMART; SM00409; Ig; 1.
CC SMART; SM00445; Link; 2.
CC PROSITE; PS00010; ASX_HYDROXYL; 1.
CC PROSITE; PS00022; EGF_1; 2.
CC PROSITE; PS01186; EGF_2; 1.
CC PROSITE; PS01187; EGF_CA; 1.
CC PROSITE; PS01241; LINK; 2.
CC PROSITE; PS00615; C-TYPE_LECTIN_1; 1.
CC PROSITE; PS00041; C-TYPE_LECTIN_2; 1.
CC Signal; Repeat; EGF-like domain; Calcium; Immunoglobulin domain;
CC Hyaluronic acid; Alternative splicing.
CC SIGNAL 1 26
CC CHAIN 27 3562
CC DOMAIN 37 136
CC DOMAIN 166 243
CC DOMAIN 264 345
CC DOMAIN 3254 3290
CC DOMAIN 3292 3328
CC DOMAIN 3341 3455
CC DOMAIN 3460 3518
CC DISULFID 44 129
CC DISULFID 171 242
CC DISULFID 195 216
CC DISULFID 269 344
CC DISULFID 293 314
CC DISULFID 3258 3269
CC DISULFID 3263 3278
CC DISULFID 3280 3289
CC DISULFID 3296 3307
CC DISULFID 3301 3316
CC DISULFID 3318 3327
CC DISULFID 3334 3345
CC DISULFID 3362 3454
CC DISULFID 3430 3446
CC DISULFID 3461 3504

```

```
FT DISULFID 3490 3517 BY SIMILARITY.
FT CARBOHYD 163 235 N-LINKED (GLCNAC. .) (POTENTIAL) .
FT CARBOHYD 235 329 N-LINKED (GLCNAC. .) (POTENTIAL) .
FT CARBOHYD 329 529 N-LINKED (GLCNAC. .) (POTENTIAL) .
FT CARBOHYD 529 709 N-LINKED (GLCNAC. .) (POTENTIAL) .
FT CARBOHYD 709 948 N-LINKED (GLCNAC. .) (POTENTIAL) .
FT CARBOHYD 948 1409 N-LINKED (GLCNAC. .) (POTENTIAL) .
FT CARBOHYD 1409 1479 N-LINKED (GLCNAC. .) (POTENTIAL) .
FT CARBOHYD 1479 1523 N-LINKED (GLCNAC. .) (POTENTIAL) .
FT CARBOHYD 1523 1530 N-LINKED (GLCNAC. .) (POTENTIAL) .
FT CARBOHYD 1530 1625 N-LINKED (GLCNAC. .) (POTENTIAL) .
FT CARBOHYD 1625 1751 N-LINKED (GLCNAC. .) (POTENTIAL) .
FT CARBOHYD 1751 1988 N-LINKED (GLCNAC. .) (POTENTIAL) .
FT CARBOHYD 1988 2088 N-LINKED (GLCNAC. .) (POTENTIAL) .
FT CARBOHYD 2088 2089 N-LINKED (GLCNAC. .) (POTENTIAL) .
FT CARBOHYD 2089 2507 N-LINKED (GLCNAC. .) (POTENTIAL) .
FT CARBOHYD 2507 2642 N-LINKED (GLCNAC. .) (POTENTIAL) .
FT CARBOHYD 2642 2679 N-LINKED (GLCNAC. .) (POTENTIAL) .
FT CARBOHYD 2679 2748 N-LINKED (GLCNAC. .) (POTENTIAL) .
FT CARBOHYD 2748 2762 N-LINKED (GLCNAC. .) (POTENTIAL) .
FT CARBOHYD 2762 3069 N-LINKED (GLCNAC. .) (POTENTIAL) .
FT CARBOHYD 3069 3194 N-LINKED (GLCNAC. .) (POTENTIAL) .
FT CARBOHYD 3194 3232 N-LINKED (GLCNAC. .) (POTENTIAL) .
FT CARBOHYD 3232 3545 N-LINKED (GLCNAC. .) (POTENTIAL) .
FT CARBOHYD 3545 485 MISSING (IN ISOFORM VI) .
FT VARSPLIC 485 1411
FT SEQUENCE 3562 AA; 388078 MW; 98C566E88C1602D2 CRC64;

Query Match 5.1%; Score 411.5; DB 1; Length 3562;
Best Local Similarity 19.7%; Pred. No. 4e-07;
Matches 397; Conservative 269; Mismatches 698; Indels 649; Gaps 94;

QY 49 TKLPQ--KNQLSGVNGVAE-QGDVHVQ-----EENQEQEEREVDVQGRES 96
DB 433 TLLPQTVDGEISPYDTLGRTEYDVSPRLTESTSAALEVENTHSEALSSE---QGRSES 489
QY 97 VREKDRVEANAANSTAV-----EDITK-----DQEQETSEIIIEIPASENNVEE 140
DB 490 TEDAFTSVVFQDSTAVAKSSTGSWEDIETGDTQKHGDQNTQIEVGPVMTATDSLVA 549
QY 141 MVQPAESQANDVGFKKFVGFVGFKFT---VKDKNEKSD---TV-----QLLTAVKXO--- 186
DB 550 SORLEPRTGSSVSLTKENLVLSHSTKEPTKKSMEAKSKLLTVTVIPKALFDQDYLTT 609
QY 187 EGGAE-----ASVAGAGHQEPV-ETAVGESAKSELKQSTEQEGT--L 230
DB 610 GEGRESMYTVPDRVSGVALVSPESDVPVAVSETLMDELAVTTG---QSSADESTPFI 666
QY 231 KQQSSTEIPLQAESDQAAEEBAKD----- 255
DB 667 KFSSTATELDNEA-SAEGRSREDLKDVHLTTSSGIPVPSFTLTANETGSEVTALSESTSAP 725
QY 256 -----EGE---EKQEKPTKSPESPSPVN-----SETTSSFK 285
DB 726 QKPEEGITSVLHSSQTEGSAILEKQBK--TKPEMSIIDAKVLVYITVVPASVTAGSEG 783
QY 286 KF-----FTHG-----WAGWRK-----KTSFKK-----SKEDDLETAE----- 313
DB 784 RFGSEKFTHTPPVSGMWLQDKQVYVTEETSHTKRIELOTEDDISGMEPTSSPGQIIIEY 843
QY 314 -KRKEQAEKVDEEK---EKTPASEQEPADTDQAR-----LSADYKVV---E 357
DB 844 TKHLGAPVSAVTDETKTSMETAETESDEEVVSADFDQTKGTTTEVFHTSSSLDLKEKFTLSK 903
QY 358 LPLEDQVGLDEASSEBKCAPLAT-----EVFEKMEAHQEVVAEVHVSVE----- 403
DB 904 IPESSATVKFSFSSSGTVPFTAATVLEVTHEADETSGYVNLNMTFTSPGEQKATE 963
QY 404 -----KTEEEQGGGGE-----AEGGVVREGTGESLPPEKLAEPQOEVAEPAEELMKS 452
DB 964 KSPATSAEDEVSTGTEISKYMTWEGGQISSVT--SAEKESVAALQOE--REEQPSVGLPET 1019
QY 453 REMCVSGGDHTQLTDLSPEEKTLPKHPGEGIVSEVEMLSQERI----- 495
```

```
DB 1020 KE-----PFKFTDVTETTPOR-EGDTSLVPTVVGSEDI GEMQVTDHTSFDSIHT 1071
QY 496 --KVQSPPLKLFSSSGLKSKQKGRGGG-GDEEPGEYQHHT-----ESPESA 545
DB 1072 EATVTSTKASEVFP---KELSTKDQDRELGTAMGSTLPVTSQVMEHQKTTAGFESPQT 1127
QY 546 DQKQGESSASSPEEPETTCLE-----KGPLEAPQ-----DGEAREGTTSDGE 588
DB 1128 TOEKIDEMGSAIDEMYPA TELSVPALMLTEYQVGSVPVETSTRSLHLTGTPKAETADGE 1187
QY 589 KK-----REGITPWASPKMVTPKKRVRRPSE----- 615
DB 1188 EKITEAVPTFGTOAKVYESKGTTRTREDRDVGSWNSVLPPTMLSSPSTAGSISLLTIG 1247
QY 616 -----SDKEELEKVKSATLS--TDST--VSEMO-----DEVKTGVEEQK--- 652
DB 1248 ASPSQTPGSGIGISEELEEVKVPFSSRATDKTIVISDLTSSISAVDKIPTASAKPFVS 1307
QY 653 -----PEEPKRVDTSVSWEALICVSSKKRARKASSDDEGGPRTLGGDSHRAEAS 705
DB 1308 SKSPRIIPE-----DEEVTSSDIIVIDESI SPS-KASAEDDL TG-----KMVEPE 1352
QY 706 KDKAECTDA-----VPASTQEQDQAQSSSPPEPAGSPSEGEVSTW----- 746
DB 1353 IDKEYFTSTATAVARPTAPPTVMEATEALQOEVSPTSHPDSGTDIRLYVITQIGNDTD 1412
QY 747 -----ESFKRLVTPRKKSKEEKARDSSVEQLSTE-IPSRSEESVVSIKKFIPIGRR 798
DB 1413 HPWNEFLDLFSRHILPHAVDEHTD--AESAQTEPCTSDSVQDSSE--YIILDPFPF--- 1465
QY 799 KRADQKQOATVDSGPGVINEEDPNVAVPLSEYNAVERKMEQAQNTLPQL----- 854
DB 1466 -NFMDFEEEDCE-----NTTDVTP--PALQFINGKQVTSAPKSTKAEARSQ 1514
QY 855 LGAVVYSELSKTLVHTVSVAVIDGTGTVSVSVEERSPSWIS--VTEPLEHTAGAM-P 911
DB 1515 IESVAHKNVTFQSINETNTFI ISETAEGTQWPSKAGEVMGAFVTPQ--TADVAMLE 1571
QY 912 PV-----EYVTEKDI IAEETPVLTQTLPEGKDA---HDDMVTSEVDFTSEAVTATSEAL 964
DB 1572 PVYSGSEVTTDKYLEITSVVEQSPKKNKVTMMHG----- 1608
QY 965 RTEEVTASGABETDMDVSAVQLDSDPTTBEATPVQEVESGLVDTBEEBQQTQAILQA 1024
DB 1609 -----TEESSTKDTKNLLITNE--SSGSGSTESOLSRSVFTTEILTMSSHED----- 1653
QY 1025 VADKVEESQVATQTVORTGSKALEKVEEVEDSEVLASEKEKDVMPKGPVQEAAGHL 1084
DB 1654 -SEKISHTTSVPTILSVERS---AVTAAPSADSDTATVGID-VKDLIPKG----- 1698
QY 1085 AQGSETGOATPESLEVPVTDVH-----VATCQVIKLQQLMEQAVAPESSETLTDSE 1138
DB 1699 -----GTATPGNYKSTIKLDAEPFENPEATSHHTKPDMTASSFVILESGDVEENS 1752
QY 1139 TNGSTPLADSDTADGTQDDTIDSDQSKATAAVROSQVTEEBEATAQKBEPTLPNNVPA 1198
DB 1753 T-----LAGAMTTE-TAVAETLSVQDT-----SLGSGTVLPTEISVTISBITPALPG--- 1798
QY 1199 QBEHGEERDVLPTQOELTAAAVPLAKTEVQGEVGDVMDLGEKVEEQEVFVHSGPN 1258
DB 1799 -----GTRILYSTFDQSEATVSNFVSELIMEQVY---GSSVATEKKY-----ED 1841
QY 1259 SOKAADVTYDSEVMGVAGQOEKE-----STEVSQLSLE-----EGEMETDVE 1300
DB 1842 EKEVQTTVSSGEI STTDAGKSELDEFQSTNVEVTVSQEPTPLREIVPTGTMMSEIK 1901
QY 1301 KEKRE--TKPEQVSEEGEQTAAPEHEGYKPVLTLDMPSSERGA-LGSLGSGSLPD 1357
DB 1902 KVTATPFLREKLFINEGSAEPPADLFAGSPTRKVVSTDSPTDSDGSDIDVITESATLFS 1961
QY 1358 QKAGCIEVQ-----VOSLDTT-----VTQTAEAV 1382
```

Db 1962 VPSRVIETQTVKHEGNIIVISVLSKNTTTEYEHIGTGPGVTSVSSGSDGTERSEVA 2021
 Qy 1383 EKVIETVVISFEGESP--ECVGHLLPAKSGATGGMWTLQHAEDTVPLGPPSOAEISPI 1440
 Db 2022 IEMSEVFEFSTENGEPQDAVPTVTPSDIKSRLG-----SREVTSHVTPVTRTDLET 2076
 Qy 1441 I-VTAPAEPTLHPD-----LGGESIASQSRSESEEDKDPADGADGKSTALEKTKAE 1493
 Db 2077 AEVTSPEVSVNNSTLDVTWHTGHTIRAVESTESKKGK-----GSFSAVSLGKILME 2129
 Qy 1494 ---PELLEESKSNKIVLN-----VIGTAVDQFARTETAPB----- 1526
 Db 2130 HOSGELKNDSSSTTKLMSNGPTEKLGSHSPFDGSGRAFTTESFTASVSPTKRPP 2189
 Qy 1527 -----THAYDSOTQVPACRLDSREPRNCWTK-----MKDAMKEH- 1560
 Db 2190 QEOYGRKTVSMPSAVVHAH-----TAEFPELVSTEHDTISLQVTVDTMEBK 2237
 Qy 1561 ----PVQPRREDLQVTVLEAMA-QPRKCLPR 1587
 Db 2238 ANELTVTSFATNLPLESDVSHMEDPREILPK 2270

RESULT 9
 PLE1 RAT STANDARD, PRT: 4687 AA.
 ID PLE1_RAT P30427; 008879; 008880; 008881;
 DC 01-APR-1993 (Rel. 25, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Plectin 1 (PLTN) (PCN).
 GN Plectin.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=glial tumor;
 RX MEDLINE=91268156; PubMed=2050743;
 RA Wiche G., Becker B., Luber K., Weitzer G., Castanon M.J.,
 RA Hauptmann R., Stratawa C., Stewart M.;
 RT "Cloning and sequencing of rat plectin indicates a 466-kD polypeptide
 chain with a three-domain structure based on a central alpha-helical
 coiled coil.";
 RT J. Cell Biol. 114:83-99(1991).
 RN [2]
 RP REVISIONS.
 RC TISSUE=glial tumor;
 RX MEDLINE=96210632; PubMed=8633055;
 RA Liu C.-G., Maercker C., Castanon M.J., Hauptmann R., Wiche G.;
 RT "Human plectin: organization of the gene, sequence analysis, and
 RT chromosome localization (8q24).";
 RL Proc. Natl. Acad. Sci. U.S.A. 93:4278-4283(1996).
 RN [3]
 RP PARTIAL SEQUENCE FROM N.A. (ISOFORMS 2; 3; 4), AND TISSUE SPECIFICITY.
 RC TISSUE=glial tumor;
 RX MEDLINE=97321050; PubMed=9177781;
 RA Elliott C.E., Becker B., Oehler S., Castanon M.J., Hauptmann R.,
 RA Wiche G.;
 RT "Plectin transcript diversity: identification and tissue distribution
 RT of variants with distinct first coding exons and rodless isoforms.";
 RL Genomics 42:115-125(1997).
 CC -I- FUNCTION: INTERLINKS INTERMEDIATE FILAMENTS WITH MICROTUBULES AND
 CC MICROFILAMENTS AND ANCHORS INTERMEDIATE FILAMENTS TO DESMOSES OR
 CC HEMIDESMOSES. MAY BE INVOLVED NOT ONLY IN THE CROSSLINKING AND
 CC STABILIZATION OF CYTOSKELETAL INTERMEDIATE FILAMENTS NETWORK, BUT
 CC ALSO IN THE REGULATION OF THEIR DYNAMICS.
 CC -I- SUBUNIT: HOMODIMER OR HOMOTETRAMER.
 CC -I- ALTERNATIVE PRODUCTS: 4 ISOFORMS; 1 (SHOWN HERE), 2, 3 AND 4; ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -I- TISSUE SPECIFICITY: WIDELY EXPRESSED WITH HIGHEST EXPRESSION IN
 CC SKELETAL MUSCLE AND LOWEST IN THYMUS.

CC -I- DOMAIN: THE N-TERMINUS INTERACTS WITH ACTIN, THE C-TERMINUS WITH
 CC VIMENTIN, DESMIN, GFAP, CYTOKERATINS, LAMIN B; WHEREAS BOTH THE N-
 CC AND THE C-TERMINUS CAN BIND INTEGRIN BETA-4.
 CC -I- PTM: PHOSPHORYLATED BY CDC2; REGULATES DISSOCIATION FROM
 CC INTERMEDIATE FILAMENTS DURING MITOSIS (BY SIMILARITY).
 CC -I- SIMILARITY: CONTAINS 1 ACTIN-BINDING DOMAIN.
 CC -I- SIMILARITY: CONTAINS 2 CALPONIN-HOMOLOGY (CH) DOMAINS.
 CC -I- SIMILARITY: CONTAINS 3 PLECTIN REPEATS.
 CC -I- SIMILARITY: CONTAINS 4 SPECTRIN REPEATS.
 CC -I- SIMILARITY: BELONGS TO THE PLEKIN OR CYTOLINKER FAMILY.
 CC -I- CAUTION: ISOFORM 4 IS A FRAGMENT AT THE N-TERMINUS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 CC EMBL; X59601; CA442169.1; -;
 CC EMBL; U96274; AAC53209.1; -;
 CC EMBL; U96275; AAC53210.1; -;
 CC EMBL; U96276; AAC53211.1; -;
 CC PIR; A39638; A39638.
 CC PIR; S21876; S21876.
 CC HSSP; Q01082; 1BR.
 CC InterPro; IPR001589; Actbind actin.
 CC InterPro; IPR001715; Calponin-like.
 CC InterPro; IPR001101; Plectin repeat.
 CC InterPro; IPR005326; S10_plectin.
 CC InterPro; IPR002017; Spectrin.
 CC Pfam; PF00307; CH; 2.
 CC Pfam; PF00681; Plectin; 21.
 CC Pfam; PF03501; S10_plectin; 1.
 CC SMART; SM00033; CH; 2.
 CC SMART; SM00250; PLEC; 33.
 CC SMART; SM00150; SPEC; 4.
 CC PROSITE; PS00019; ACTININ_1; FALSE NEG.
 CC PROSITE; PS00020; ACTININ_2; FALSE NEG.
 CC PROSITE; PS50021; CH; 2.
 CC Coiled coil; Repeat; Structural protein; Cytoskeleton; Actin-binding;
 CC KW Phosphorylation; Alternating splicing.
 CC DOMAIN 1 1473
 CC FT 1474 2758 GLOBULAR 1.
 CC FT 2759 4687 GLOBULAR 2.
 CC FT 4688 406 ACTIN-BINDING.
 CC FT 407 288 CH 1.
 CC FT 301 403 CH 2.
 CC FT 404 722 SPECTRIN 1.
 CC FT 723 827 SPECTRIN 2.
 CC FT 828 933 SPECTRIN 3.
 CC FT 934 1418 SPECTRIN 4.
 CC FT 1419 1692 COILED COIL (POTENTIAL).
 CC FT 1693 2760 COILED COIL (POTENTIAL).
 CC FT 2761 2828 PLECTIN 1.
 CC FT 2829 2866 PLECTIN 2.
 CC FT 2867 2904 PLECTIN 3.
 CC FT 2905 2942 PLECTIN 4.
 CC FT 2943 2980 PLECTIN 5.
 CC FT 2981 3018 PLECTIN 6.
 CC FT 3019 3156 PLECTIN 7.
 CC FT 3157 3194 PLECTIN 8.
 CC FT 3195 3232 PLECTIN 9.
 CC FT 3233 3270 PLECTIN 10.
 CC FT 3271 3308 PLECTIN 11.
 CC FT 3309 3346 PLECTIN 12.
 CC FT 3347 3385 PLECTIN 13.
 CC FT 3386 3424 PLECTIN 14.
 CC FT 3425 3463 PLECTIN 15.
 CC FT 3464 3501 PLECTIN 16.
 CC FT 3502 3540 PLECTIN 17.
 CC FT 3541 3579 PLECTIN 18.

DB 2805 SNEELRLQGHHTVAELTQREDEVRHYLKGSSAGLLKPTNEKLSVTYALQRLSLP- 2863

OY 1400 CVAHLLPFAKSSATG-----GHWTLOHAEIVPLGPE 1432

DB 2864 --GTALLLEAQAASGFLDPVARNRLTYNEAVKEGVGPE 2902

RESULT 10

PLEI CRIGR STANDARD; PRT: 4473 AA.

AC 090155;

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Plectin 1 (PLTN) (PCN) (300-kDa intermediate filament-associated protein) (IFAP300) (Fragment).

GN Plectin.

OS Cricetulus griseus (Chinese hamster).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae; Cricetulus.

OX NCBI_TaxID=10029;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=20334248; PubMed=10873583;

RA Clubb B.H., Chou Y.-H., Herrmann H., Svickina T.M., Borisy G.G., Goldman R.D.;

RT "The 300-kDa intermediate filament-associated protein (IFAP300) is a hamster plectin ortholog.";

RL Biochem. Biophys. Res. Commun. 273:183-187(2000).

CC [2]

CC PHOSPHORYLATION.

CC MEDLINE=96215219; PubMed=8626512;

CC Malecz N., Foisner R., Stadler C., Miche G.;

CC "Identification of plectin as a substrate of p34cdc2 kinase and mapping of a single phosphorylation site.";

CC J. Biol. Chem. 271:8203-8208(1996).

CC [1]

CC FUNCTION: INTERLINKS INTERMEDIATE FILAMENTS WITH MICROTUBULES AND MICROFILAMENTS AND ANCHORS INTERMEDIATE FILAMENTS TO DESMOSOMES OR HEMIDESMOSOMES. MAY BE INVOLVED NOT ONLY IN THE CROSSLINKING AND STABILIZATION OF CYTOSKELETAL INTERMEDIATE FILAMENTS NETWORK, BUT ALSO IN THE REGULATION OF THEIR DYNAMICS.

CC [1]

CC SUBUNIT: HOMODIMER OR HOMOTETRAMER.

CC [1]

CC DOMAIN: THE N-TERMINUS INTERACTS WITH ACTIN. THE C-TERMINUS WITH VIMENTIN, DESMIN, GAP, CYTOKERATINS, LAMIN B; WHEREAS BOTH THE N- AND THE C-TERMINUS CAN BIND INTEGRIN BETA-4.

CC [1]

CC PTM: PHOSPHORYLATED BY CDC2; REGULATES DISSOCIATION FROM INTERMEDIATE FILAMENTS DURING MITOSIS.

CC [1]

CC SIMILARITY: CONTAINS 1 ACTIN-BINDING DOMAIN.

CC [1]

CC SIMILARITY: CONTAINS 2 CALPONTIN-HOMOLOGY (CH) DOMAINS.

CC [1]

CC SIMILARITY: CONTAINS 32 PLECTIN REPEATS.

CC [1]

CC SIMILARITY: CONTAINS 4 SPECTRIN REPEATS.

CC [1]

CC SIMILARITY: BELONGS TO THE PLAKIN OR CYTOLINKER FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

CC EMBL; AF260753; AAF70372.1; -.

CC HSSP; O01082; 1BRK.

DR InterPro; IPR001589; Actbind actnin.

DR InterPro; IPR001715; Calponton-like.

DR InterPro; IPR001101; Plectin_repeat.

DR Pfam; PF00307; CH; 2.

DR Pfam; PF00681; Plectin; 20.

DR SMART; SM00033; CH; 2.

DR SMART; SM00250; PLEC; 32.

DR SMART; SM00150; SPEC; 4.

DR PROSITE; PS00019; ACTININ_1; PARTIAL.

DR PROSITE; PS00020; ACTININ_2; FALSE NEG.

DR PROSITE; PS50021; CH; 2.

KM Coiled coil; Repeat; Structural protein; Cytoskeleton; Actin-binding; Phosphorylation.

FT NON TER 1 1

FT DOMAIN <1 1259

FT DOMAIN 1260 2544

FT DOMAIN 2545 4473

FT DOMAIN <1 192

FT DOMAIN <1 74

FT DOMAIN 87 189

FT DOMAIN 449 508

FT REPEAT 529 613

FT REPEAT 626 719

FT REPEAT 1104 1204

FT DOMAIN 1258 2548

FT REPEAT 2615 2652

FT REPEAT 2653 2690

FT REPEAT 2691 2728

FT REPEAT 2729 2766

FT REPEAT 2770 2804

FT REPEAT 2905 2942

FT REPEAT 2943 2980

FT REPEAT 2981 3018

FT REPEAT 3019 3056

FT REPEAT 3057 3094

FT REPEAT 3274 3311

FT REPEAT 3312 3349

FT REPEAT 3350 3387

FT REPEAT 3388 3425

FT REPEAT 3429 3463

FT REPEAT 3609 3646

FT REPEAT 3647 3684

FT REPEAT 3685 3722

FT REPEAT 3723 3760

FT REPEAT 3764 3797

FT REPEAT 3800 3834

FT REPEAT 3852 3889

FT REPEAT 3927 3964

FT REPEAT 3965 4003

FT REPEAT 4007 4041

FT REPEAT 4043 4094

FT REPEAT 4197 4234

FT REPEAT 4235 4272

FT REPEAT 4273 4310

FT REPEAT 4311 4348

FT REPEAT 4349 4386

FT DOMAIN 4039 4089

FT DOMAIN 4414 4429

FT MOD RES 4328 4328

SO SEQUENCE 4473 AA; 509015 MW; E144615D361E3484 CRC64;

Query Match 5.0%; Score 402.5; DB 1; Length 4473;

Best Local Similarity 19.2%; Pred. No. 1e-06;

Matches 343; Conservative 280; Mismatches 774; Indels 389; Gaps 64;

OY 8 EGRSPRQPAAGSP-----TSELVLSGHGPAFAEASGAPADADPATKLPQK-----NQO 57

DB 1273 EERLAEQQAERERERLAGEALEEKOROLAENHQAQAQAELE -AQELQRRQGEVARRE 1331

OY 58 LSSVNGVAEQGDVHVOENQ---EGQEEVDE---DVQRESEDEVEKDRVEEMAANS 110

DB 1332 EAANV--AQQQRKRSIQBELQHRLQSSFAEIQAKAQVQEALESRRIRIEIRVVALQLET 1389

OY 111 TAVEDITDGGQGETSET-IEQIPASENVVEVQPAESQANDVGFKVFKFVGFKFTYK- 168

DB 1390 TERQSGAGAGELQALPARAEFAEAQKQAQAEFAERLRQVDQESQRKQAQEAELARVYA 1449

OY 169 --KDKNEKSDTYQLLVTKKDBEGEAGASVAGADHQ-----PPSVETANGESAKSES 217

Db 1450 QAEAAQKQALQALELRLQAEAEARRLRQAQARARQVQVALETAQRAEAEVELQSKRA 1509
Qy 218 ELKOSTEQEGTLKQEQ-SSTEIPLOAE---SDQAAEEEAQDEGEKQEKPTKSPSPS 273
Db 1510 SFAEKTAQLETLQEEHVTVTLQREAAQQAQAEAREARELERWQLKANEALR 1569
Qy 274 SPVNETTSFKFFTHGWAGWRKTSFKSKEDDLTETAKRKEQAEKVDESEKTEP 333
Db 1570 LRQAEVAQOQSL-----AQADAQKQEEAEAREARRRGAEEQAVRQRELAQE- 1619
Qy 334 ASBQEPADTDQARLSADYKVELPLEDOVGDLAS--SEKCAPL---ATEVPDEKMEA 309
Db 1620 LEKQRLAEGTAQORLAEOELIRLRAETEGQEQORQLLEELARLOREATAATHKQEL 1679
Qy 390 HQEVV-----AEHVSTVETKEBEGGGGE-----AEGVVVEGTGESL 428
Db 1680 EAEALKVAEMEVLLASKARAEBSRSTSEKSKORLEAEADRPRELAEEAARLALAEA 1739
Qy 429 PPEK-LAPQOVPOEAPAPAEELMKSRMCVSGGDHTOLTDLSPBEKTLPHPEGIVSEVE 487
Db 1740 KRQRLAEEADAARQRAE-ABRVLTKELAIS-----BATRLKTEABTALKEKEAENLR 1793
Qy 488 MLSSQERIKVQSGPLKGLFSSGGLKLGKKGKGGGDEBEGVQHIHTSPESADE 547
Db 1794 RLAEDEAFQ-----RRLEQALHKAADIEERLAQ-----LRKASESELER 1834
Qy 548 QKG-ESSASPPEPETTCLEKGPLAPQDGEAE-----EGTSDGKKREGITPW 597
Db 1835 QKGLVEDTLQRQVBEILALKVSEFKAAGAAELELGRINSSAEDTMRSEKQAEQE 1894
Qy 598 ASFKQMTPKVRVRPSBK-----EELKVKQATLSSTSDTSTVSEMODEVKTVGBE 650
Db 1895 AARQRLAEEQORREAREVQKSLAAEEBAARQKAALEEVERLKAV-EEARRLRER 1953
Qy 651 QKPEEPK-----RRVDTSVSWALICVGSKKRKAASSDDGEGPRTLGGDSHR 700
Db 1954 AEQSAQOLQAEAAQKRLQAEKAHAFV--QORREELQOTLQOQSMLERLURGAEEA 2011
Qy 701 AEPASKD-KEAGTDVAPSTQEQDQAQGS-----SSPEPAGSPSEGGVSTWSEFKRLVT 754
Db 2012 ARRAAEAEAEQAEAREAAQSRKQVEAEERLQSGAEQAAQAAQAAAA-----2061
Qy 755 PRKSKSLEKEAEDSVQSLSTEIPSRSESWISIKKFIPIGRKKRAGDKQEQATVEDS 814
Db 2062 -EKLKEAEQAEARRAQAQEAALQKQADAEMAEKHKFAEQTLRQAEVQELTTLR-- 2118
Qy 815 GPVEINDDNPVAVPLSEVNAVERKME-AQNTLQPLLGAIVYS-EELSKTLVHTV 872
Db 2119 --LQLEETD-HQKSILD-EELQRLKAEVTEAARQSRQVEBELFSVRVOMBELGK-----2168
Qy 873 SVAVIDGTRAVTSVEERSPSWISASVTEP-LEHTAGAMPPEVTEVKDIIABETPVLQ 931
Db 2169 -----LKARIEAENRALILROKNTQRFLEBAERKQVAAEAAARLSVAQAQEAARLQ 2221
Qy 932 -----TLPEGKADHDMVTSEVDFTSVAVTETSEALRTEBEVTASGAETTMVSAVSQ 987
Db 2222 LAEBDLAQORALAEKMLKEMQAVQAEATRLKAEAEELQQOKELAQEQARRLQ-----2274
Qy 988 LTSDPDTTEATPVQEVESGLVDFTEERQTOAILQAVADVKV-----EES 1033
Db 2275 -----DKEQNAQQLVEETQGFORTLEVRQRLQEMSAERLKLURMAEMRAQARAEDA 2329
Qy 1034 Q-----VPATQT--VORTGS-----KALEKVEBEVEDSE-- 1060
Db 2330 QRFRKQAEIEGKHLRTELATQEKVTLVQTLQORQSDHDAERLRAIAELEREKEKLK 2389
Qy 1061 -----VLAEEKQKNDMPKGPVQAGAEHLAQSGSETQCATPESLEVEPVDVHDVATQC 1114
Db 2390 QEAKLLQKSEEMTVQEQEQLQTOALQKSFLEKSDLSLLQRERFIEQEKAKLEQLFQDE 2449
Qy 1115 VIKLOQLME-----QAVAPSETLTDSETNGSTPLADSDTADGTQDE-----TIDSQDS 1165
Db 2450 VAKAQQLREBQORQORQMEQEKQELVASMEBAARRRQCEAEAEVRRKQOEELQHLQELQOQ 2509

Qy 1166 KATAAVRQSOVTE-----EEATAQKESFSTLPNNVPAQEEHGEPEGRD 1209
Db 2510 EKLLEENQRLRQLRLEEBEHAALAHSEETAATAQAAAKALPN-----GRD 2557
Qy 1210 VLEPTQOEL-----TAAAVPLAKTEVQGEVDMLDGKVKKEQEVFVHSGNSOK 1261
Db 2558 ALDGPSEVEPEHAFFGLRQKVPALQRLQREAG-----ILSAEELQRLQOCHTTVAELSQR 2611
Qy 1262 AADVTYDSEVMGVAGQCEKSETEVQS-----LSLEEGEMET-----DVEKE 1302
Db 2612 EDVRYLQGRSSIFAGLLPTEDEKLSVYTALQRLQLSPGTALILLEAQAASGFLDPVRN 2671
Qy 1303 KRETKPEQVSEEGEQETAAPHEHGYGKPVLTLDMPSSR---GKALGSLGGSPSLPDQD 1359
Db 2672 RRLTVNE-----PVKEGVVG-PELHHKLLSAERAVTGYKDPYVTEQISLFOAM 2718
Qy 1360 KACCEVQVQSLDVTVTQTAEAVEKVIETV-----VISETG-- 1395
Db 2719 KK---DLLVRDHAIRLLEAQIATGGIITDVHSHRVPDVVAYQRGYFDEEMSRILADPGDD 2775
Qy 1396 -----ESPECVGAHLLPAEKSSATGCH--WTLQHAEDTVPLG 1430
Db 2776 TKGFFDPNTHENLYLQLLERCVEDEPE-TGLHLLPLTDKAAGGELVYDTTEARDVF--- 2831
Qy 1431 PESQAESIPITVPAPESTLHPDLOGE-ISASQSRSESEEDKPDAGDADGKESTAIEKV 1489
Db 2832 -EKATVSAFFGFGQKTVTIWEIINSEYFTAEOQRDLRLQFRGT-----RITVEKI 2881
Qy 1490 LKAEPEILELESKNKIVLNVTQAVDQPARTETAPETTHAYDSOTQ 1535
Db 2882 IKIVITVBEQERKQGLCFEGLRALVPAAELLESVISHELYQOLQ 2927
RESULT 11
PLE1_HUMAN
ID PLE1_HUMAN STANDARD; PRT; 4684 AA.
AC Q15149; Q16640; Q15148;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Plectin 1 (PLN) (PCN) (Hemidesmosomal protein 1) (HD1).
GN PLE1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Placenta;
RX MEDLINE=96210632; PubMed=8633055;
RA Liu C.-G., Maercker C., Castranon M.J., Hauptmann R., Wiche G.;
RT "Human plectin: organization of the gene, sequence analysis, and
RT chromosome localization (8q24).";
RL Proc. Natl. Acad. Sci. U.S.A. 93:4278-4283 (1996).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 2 AND 3), AND DISEASE.
RX MEDLINE=96312447; PubMed=8698233;
RA McLean W.H.I., Pulkkinen L., Smith F.J.D., Rugg E.L., Lane E.B.,
RA Bullrich F., Burgeson R.B., Amano S., Hudson D.D., Owarike K.,
RA McGrath J.A., McMillan J.R., Eady R.A.J., Leigh I.M., Christiano A.M.,
RA Uitto J.;
RT "Loss of plectin causes epidermolysis bullosa with muscular dystrophy:
RT cDNA cloning and genomic organization.";
RL Genes Dev. 10:1724-1735 (1996).
RN [3]
RP VARIANT MD-EBS 1003-GLN-ALA-1005 DEL.
RX MEDLINE=97049959; PubMed=8894687;
RA Pulkkinen L., Smith F.J.D., Shimizu H., Murata S., Yavita H.,
RA Hachisuka H., Nishikawa T., McLean W.H.I., Uitto J.;
RT "Homozygous deletion mutations in the plectin gene (PLEC1) in patients
RT with epidermolysis bullosa simplex associated with late-onset
RT muscular dystrophy.";

RL Hum. Mol. Genet. 5:1539-1546(1996).
 RN [4]
 RP VARIANT MD-EBS LEU-429 INS.
 RX MEDLINE=21090821; PubMed=11159198;
 RA Bauer J.W., Rouan F., Kofler B., Reznicek G.A., Kornacker I.,
 RA Mus W., Hametner R., Klausegger A., Huber A., Pohl-Gabo G.,
 RA Wiche G., Uitto J., Hinner H.
 RT "A compound heterozygous one amino-acid insertion/nonsense mutation in
 RT the plectin gene causes epidermolysis bullosa simplex with plectin
 RT deficiency.";
 RL Am. J. Pathol. 158:617-625(2001).
 CC -1- FUNCTION: INTERLINKS INTERMEDIATE FILAMENTS WITH MICROTUBULES AND
 CC MICROFILAMENTS AND ANCHORS INTERMEDIATE FILAMENTS TO DESMOSES OR
 CC HEMIDESMOSES. COULD ALSO BIND MUSCLE PROTEINS SUCH AS ACTIN TO
 CC MEMBRANE COMPLEXES IN MUSCLE. MAY BE INVOLVED NOT ONLY IN THE
 CC CROSSLINKING AND STABILIZATION OF CYTOSKELETAL INTERMEDIATE
 CC FILAMENTS NETWORK, BUT ALSO IN THE REGULATION OF THEIR DYNAMICS.
 CC -1- SUBUNIT: HOMODIMER OR HOMOTETRAMER.
 CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; 1 (SHOWN HERE), 2 AND 3; ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED WITH HIGHEST LEVELS IN
 CC MUSCLE, HEART, PLACENTA AND SPINAL CORD.
 CC -1- DOMAIN: THE N-TERMINUS INTERACTS WITH ACTIN, THE C-TERMINUS WITH
 CC VIMENTIN, DESMIN, GAP, CYTOKERATINS, LAMIN B; WHEREAS BOTH THE N-
 CC AND THE C-TERMINUS CAN BIND INTEGRIN BETA-4.
 CC -1- PIN: PHOSPHORYLATED BY CDC2, REGULATES DISSOCIATION FROM
 CC INTERMEDIATE FILAMENTS DURING MITOSIS (BY SIMILARITY).
 CC -1- DISEASE: DEFECTS IN PLECI ARE THE CAUSE OF EPIDERMOLYSIS BULLOSA
 CC WITH MUSCULAR DYSTROPHY (MD-EBS OR EB-MD), AN AUTOSOMAL RECESSIVE
 CC DISORDER CHARACTERIZED BY EPIDERMAL BLISTER FORMATION AT THE LEVEL
 CC OF THE HEMIDESMOSE AND ASSOCIATED LATE-ONSET MUSCULAR DYSTROPHY.
 CC -1- SIMILARITY: CONTAINS 1 ACTIN-BINDING DOMAIN.
 CC -1- SIMILARITY: CONTAINS 2 CALPONIN-HOMOLOGY (CH) DOMAINS.
 CC -1- SIMILARITY: CONTAINS 3 PLECTIN REPEATS.
 CC -1- SIMILARITY: CONTAINS 4 SPECTRIN REPEATS.
 CC -1- SIMILARITY: BELONGS TO THE PLAKIN OR CYTOLINKER FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U534367; CAA91196.1; -
 DR EMBL; U53204; AAB05427.1; -
 DR EMBL; U63610; AAB05428.1; -
 DR EMBL; U63609; AAB05428.1; JOINED.
 DR EMBL; X97053; CAA65765.1; -
 DR HSSP; Q01082; IBKR.
 DR Genew; HGNC:9069; PLEC1.
 DR MIM; 601282; -
 DR MIM; 226670; -
 DR InterPro; IPR001589; Actbind actinin.
 DR InterPro; IPR001715; Calponin-like.
 DR InterPro; IPR001101; Plectin_repeat.
 DR InterPro; IPR005326; S10_plectin.
 DR InterPro; IPR002017; Spectrin.
 DR Pfam; PF00307; CH; 2.
 DR Pfam; PF00681; Plectin; 19.
 DR Pfam; PF03501; S10_plectin; 1.
 DR SMART; SM00033; CH; 2.
 DR SMART; SM00250; PLEC; 33.
 DR SMART; SM00150; SPEC; 5.
 DR PROSITE; PS00019; ACTININ_1; FALSE NEG.
 DR PROSITE; PS00020; ACTININ_2; FALSE NEG.
 DR PROSITE; PS50021; CH; 2.
 KW Coiled coil; Repeat; Structural protein; Cytoskeleton; Actin-binding;
 KW Phosphorylation; Alternative splicing; Epidermolysis bullosa;
 KW Disease mutation.
 FT DOMAIN 1 1470 GLOBULAR 1.
 FT 1471 2755 CENTRAL FIBROUS ROD DOMAIN.

FT DOMAIN 2756 4684 GLOBULAR 2.
 FT 175 400 ACTIN-BINDING.
 FT 179 282 CH 1.
 FT 295 397 CH 2.
 FT 645 710 SPECTRIN 1.
 FT 740 824 SPECTRIN 2.
 FT 837 930 SPECTRIN 3.
 FT 1315 1415 SPECTRIN 4.
 FT 1469 2756 COILED COIL (POTENTIAL).
 FT 2826 2863 PLECTIN 1.
 FT 2864 2901 PLECTIN 2.
 FT 2902 2939 PLECTIN 3.
 FT 2940 2977 PLECTIN 4.
 FT 2981 3015 PLECTIN 5.
 FT 3116 3153 PLECTIN 6.
 FT 3154 3191 PLECTIN 7.
 FT 3192 3229 PLECTIN 8.
 FT 3230 3267 PLECTIN 9.
 FT 3268 3305 PLECTIN 10.
 FT 3306 3343 PLECTIN 11.
 FT 3385 3422 PLECTIN 12.
 FT 3523 3560 PLECTIN 13.
 FT 3561 3598 PLECTIN 14.
 FT 3599 3636 PLECTIN 15.
 FT 3640 3674 PLECTIN 16.
 FT 3820 3857 PLECTIN 17.
 FT 3858 3895 PLECTIN 18.
 FT 3896 3933 PLECTIN 19.
 FT 3934 3971 PLECTIN 20.
 FT 3975 4008 PLECTIN 21.
 FT 4063 4100 PLECTIN 22.
 FT 4101 4138 PLECTIN 23.
 FT 4139 4176 PLECTIN 24.
 FT 4177 4214 PLECTIN 25.
 FT 4218 4252 PLECTIN 26.
 FT 4265 4305 PLECTIN 27.
 FT 4319 4356 PLECTIN 28.
 FT 4408 4445 PLECTIN 29.
 FT 4446 4483 PLECTIN 30.
 FT 4484 4521 PLECTIN 31.
 FT 4522 4559 PLECTIN 32.
 FT 4560 4597 PLECTIN 33.
 FT 4250 4300 BINDING TO INTERMEDIATE FILAMENTS (BY
 FT SIMILARITY).
 FT 4 X 4 AA TANDEN REPEATS OF G-S-R-X.
 FT PHOSPHORYLATION (BY CDC2) (BY
 FT SIMILARITY).
 FT MVAAGLMRPDQLRAIYELFRFEGVAVAKDRRPSLHPHP
 FT GVTNLQWRAMASRARGVETPAMCHFPYVLTNEGIAHL
 FT ROYLALPREIYASLQRYRPVAMVMPARTPHYQAVGCP
 FT GSPPRGDLPTREORLTKRKELEVSPTVPVATTQRTIA
 FT RRPPEPAVAT -> MSGEDAVRAVSBDVSGSSGSPSGD
 FT TLPMNLKGTQSRSGGAGSGSVLDPAAERAVIRIA (IN
 FT ISOFORM 2 AND ISOFORM 3).
 FT MISSING (IN ISOFORM 3).
 FT L -> IL (IN MD-EBS).
 FT /Pftid=VAR_011336.
 FT MISSING (IN MD-EBS).
 FT /Pftid=VAR_011337.
 FT S -> F (IN REF. 2).
 FT N -> D (IN REF. 2).
 FT H -> N (IN REF. 2).
 FT I -> V (IN REF. 2).
 FT Q -> R (IN REF. 2).
 FT N -> Y (IN REF. 2).
 FT V -> A (IN REF. 2).
 FT V -> L (IN REF. 2).
 FT L -> V (IN REF. 2).
 FT V -> L (IN REF. 2).
 FT I -> M (IN REF. 2).
 FT A -> T (IN REF. 2).
 FT WLC -> RLR (IN REF. 2).
 FT Q -> E (IN REF. 2).
 FT 1 174 VARSPLIC
 FT 4625 4640 DOMAIN
 FT 4539 4539 MOD_RES
 FT 409 412 VARSPLIC
 FT 429 429 VARIANT
 FT 1003 1005 VARIANT
 FT 185 185 CONFLICT
 FT 259 259 CONFLICT
 FT 550 550 CONFLICT
 FT 560 560 CONFLICT
 FT 706 706 CONFLICT
 FT 886 886 CONFLICT
 FT 1002 1002 CONFLICT
 FT 1309 1309 CONFLICT
 FT 1321 1321 CONFLICT
 FT 1334 1334 CONFLICT
 FT 1534 1534 CONFLICT
 FT 1662 1662 CONFLICT
 FT 1688 1690 CONFLICT
 FT 1767 1767 CONFLICT

```

FT CONFLICT 1789 1789 L -> A (IN REF. 2).
FT CONFLICT 1910 1910 K -> R (IN REF. 2).
FT CONFLICT 2154 2154 K -> N (IN REF. 2).
FT CONFLICT 2160 2160 S -> Q (IN REF. 2).
FT CONFLICT 2215 2215 R -> Q (IN REF. 2).
FT CONFLICT 2244 2244 S -> A (IN REF. 2).
FT CONFLICT 3027 3027 K -> E (IN REF. 2).
FT CONFLICT 3310 3310 E -> A (IN REF. 2).
FT CONFLICT 3361 3361 F -> L (IN REF. 2).
FT CONFLICT 3408 3408 F -> L (IN REF. 2).
FT CONFLICT 3447 3447 G -> A (IN REF. 2).
FT CONFLICT 3531 3531 S -> A (IN REF. 2).
FT CONFLICT 3580 3580 R -> S (IN REF. 2).
FT CONFLICT 3589 3589 K -> Q (IN REF. 2).
FT CONFLICT 3596 3596 E -> Q (IN REF. 2).
FT CONFLICT 3616 3616 N -> H (IN REF. 2).
FT CONFLICT 3686 3686 V -> A (IN REF. 2).

Query Match
Best Local Similarity 5.0%; Score 400.5; DB 1; Length 4684;
Matches 313; Conservative 275; Mismatches 667; Indels 331; Gaps 62;

QY 8 EORSPEOPAGSD---TPSELVLSGHGPAABAGAGDADADPATKLPQK-----NGQ 57
DB 1484 EERLAEQORAEERLERLAEVAALAEKQRLAEAHQAQAQARE- AKELQORIIEVVVRE 1542

QY 58 LSSVNGVAEGDVHVRQENEGQEEVYVDEVDGORESE-----DVREKDRVREMAANS 110
DB 1543 EAAVD--AQQKRSIQBELQQLQSSAEIQAQQAQAEAAERSRLRIIEEIRVRLQLEA 1600

QY 111 TAVEDITKQOEETSEI- IQOIPASENNVEVMQPBASQANDVGFKVFGVQKFTVKK 169
DB 1601 TERQORGAEGELQALRAAEAEQAQRQAQAEERLRQVQDESQRKQRAEVELASRKA 1660

QY 170 D---KNEKSDTVOLLTVKKDGGCAE-----ASVGADHGPSPVET-----AVGESAKES 217
DB 1661 EAAAREKQALQALELRLLQAEAEERWLQAEVERARQVQVALETHQRSABAELOSRA 1720

QY 218 ELQKSTEQBGTLLKQEO- SSTEIPLQAE-----SDQAEEAEKDEGEKQEKTKSPSPS 273
DB 1721 SFAEKTAQLERSLQEEHVAQALREAEARRAQQAQAEARAEARERQLERWQLKANEALR 1780

QY 274 SPVNSETTSFKFFTHGAGWKTKTSFKSKEDDLTAERKKEQAEKVDBEEKETEP 333
DB 1781 LRLQAEVLQOKSL-----AQAEAKQKEAEAEARRRGAERQVQRELAEQE- 1830

QY 334 ASSEOPEADTDQARLSADYKVELPLEDOVGDLEAS- SEEKCAPL---ATEVFDEKWEA 369
DB 1831 LEKORQLAEGTAQORLAQAEQELINLRATEQGGQORQLLEELARLOREAAATQKKQEL 1890

QY 390 HQEVW-----AEVHVSTVKTETEEQGGG-----EAEGVVVEGTGE-----SLPPE- 431
DB 1891 EAEALKVRAEMEVILASKAKAABESRSTSEKSKORLEAEAGRFRELAEEAARLALAEAA 1950

QY 432 ----KLAEPOEVQAEAPAEELMKSRMCVSGDHTQTLTDLSPEEKTLPHKPGEIVSEVE 487
DB 1951 KRQQLAEEDAARQRAE-AERVLAEKLAAL- -GEATRL-----KT-----EAEIALKE 1995

QY 488 MLSSQRIKVGSPGLKLFSSGLKLSGKKQKRGKGGGDEBPGEYQHITHTESPESADE 547
DB 1996 KEAENER-----LRLAEDEAPQRRLEEBEQAQKHADIEERLAQ---LRKASDSLEL 2045

QY 548 QKG- -BSSASSPEPETTCLEKGPLAPQDGEAE-----EGTTSDEKKKREGITPW 597
DB 2046 QKGLVEDTLQRQVQVEBEILLALKASFKAANGAKAELELELGRISNAEDTLRSKEQAELE 2105

QY 598 ASFKQMTVTKKVRPSPESDK-----EBELEKVSATLSSTDSTVSEMQDEVKTVGEE 650
DB 2106 AARQQLAAEEERRRRAEERVQKSLAAEEBAAERQKAALEEVERLKAHV-EEARSURER 2164

QY 651 QKPEEPK-----RRVDTSVSWFALICVGSKKRKAASSDDGEGPRLTGGDSHR 700
DB 2165 AEQESAKQLQAEAAQRLQAEKAAHAFV--QQKEQELQOTLQOEQSQSVLDRLRGEAA 2222

```

RESULT 12

PGCV_HUMAN

ID PGCV_HUMAN

AC P13611; P20754; Q9UNW5; Q13010; Q13189; Q15123;

DT 01-JAN-1990 (Rel. 13, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Versican core protein precursor (large fibroblast proteoglycan)

DE (Chondroitin sulfate proteoglycan core protein 2) (PG-M) (Glia)

DE hyaluronate-binding protein) (GHP).

GN CSPG2.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

```

QY 701 ABEASKD-KEAGTDAVPASTQEQDQAQGS-----SSPEPAGSPSEGEVSTWESFKRLVT 754
DB 2223 ARRAAEAEAEARVQAEAREAAQRRQVEEAERLKQSAEQQAQAQAAA-----2272
QY 755 PRKKSXSKLEEKAEDSSVQLSTEIPSRSEESWVSFKKTPGRKKRKGQKQOATVEDS 814
DB 2273 -EKLKAEQEAARRAQAQAALRQQAADAEMEKHKFAEQTLRQKQAVEQLTLR-- 2329
QY 815 GPVEINDDPNPVPVPLSEYNAVERE---KME-----AQNTLPLQLGAVVVS-EELS 865
DB 2330 --LQLEETD-----HQKLLDEELQRLKAAEATAAARQSRQSVBEELFSVRVQMEELS 2378
QY 866 KTLVHTVSVAVIDGTRAVTSVEERSPSWISASVTEPLEHTAGAMPVPV-EEVTEKIIAE 924
DB 2379 K-----LKAITEAENRALILRDKONTORFLOEAEKMKQVAAEAARLSVAAQ 2425
QY 925 ETPLVLTQ---TLPEGKDAHDDMVTSEVDFTSFVATATETSEALRTEEVTEASGAETTD 980
DB 2426 EAARLQLAEEDLAQORALAEKMLKEMQAVQATLKAELQLQKQKELAQEQARRLQE 2485
QY 981 MSAVSQSLTSDSPDTTBEATPVOEVESGVLDTBEEERQTQAILQAVADKVK-----EESQ 1034
DB 2486 -----DKEQMAQQLAEBETQGFQRTLEAERQRLQLEMSAEARLKLKRVAMSRQA 2533
QY 1035 VPATQTVQRTGSKALEKVEVEEDSEVLASEKEKQWMPKGPVQEAAGHQAQSGTQAT 1094
DB 2534 ARAEEDAQRFRQAEBEIGEKLRHTE--LATQEKVTLVQTLQIQRQOSDHDAE-----2583
QY 1095 PESLEVPVETADVH--VATCQVILQOOL-MEQAVAPESSETLTDSETNGSTPLADSDTA 1151
DB 2584 ----RLREAIAELEREKEKLQOEAKLLQKSEEMQTVQOEFQLQOETQALQOQSFLEKDSL 2639
QY 1152 DGTQODETTDSODSKATAAVRQSOVTEEEAATAQKEEPSTLPNNVPAQOEHEGSEPGRDVL 1211
DB 2640 --LQREFIEQEKAL-----EQLFQDEVAKAQ-----LREE--QORQQOQM 2678
QY 1212 EPTQQBELTAAAVPVLAKTEVGOGEVDWLDGEKVKEEVEVHSGPNQSKAADVTYDSEV 1271
DB 2679 EQERQRLVASMEEARRRQHEAEG-----VRRKQE-----2708
QY 1272 MGVAQCQEKESTEVQSLSLEEGEMETDVEKEKETXPE-QVSEEGEQETAAPHEGTGK 1330
DB 2709 -----ELOQLEQRRQOEELLAENQRLQELQLEEQHRAALAHSEVTSAQ 2756
QY 1331 PVLTLDMPS-----SERKALGSLGSPSPDPQDKAGCIEV-QVQSL---DTTV 1375
DB 2757 VAATKTLPNGRDALDGPAAEAPEHSPFDGLRRKVSQRLQEAQILSAEELQRLAQGHVV 2816
QY 1376 TQTAEEVE-----KVIETVVISSETGE-----SPECVGAHLPLPAEKSSAT 1414
DB 2817 DELARREDVRHVLYQGRSSIAGLLLKATNEKLSVYAAALQRLQLSP---GTALILLEAQAAS 2873
QY 1415 G-----GHWTQLHAEDTVPLGPE 1432
DB 2874 GFLLDPVRNRLTVAEAVKEGVVGPE 2899

```

RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM V0).
 RX MEDLINE=95105188; PubMed=7528742;
 RA Nao M.F., Zimmermann D.R., Iozzo R.V.,
 RT "Characterization of the complete genomic structure of the human
 RL versican gene and functional analysis of its promoter.";
 RN J. Biol. Chem. 269:32999-33008(1994).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM V1).
 RX TISSUE=Placenta;
 RL MEDLINE=90053982; PubMed=2583089;
 RA Zimmermann D.R., Ruoslahti E.,
 RT "Multiple domains of the large fibroblast proteoglycan, versican.";
 RL EMBO J. 8:2975-2981(1989).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM V2).
 RX TISSUE=Glial tumor;
 RL MEDLINE=95105187; PubMed=7806529;
 RA Doures-Zimmermann M.T., Zimmermann D.R.,
 RT "A novel glycosaminoglycan attachment domain identified in two
 RL alternative splice variants of human versican.";
 RL J. Biol. Chem. 269:32992-32996(1994).
 RN [4]
 RP SEQUENCE OF 2711-3396 FROM N.A.
 RX TISSUE=Lung fibroblast;
 RL MEDLINE=88007514; PubMed=2820964;
 RA Krusius T., Gehlsen K.R., Ruoslahti E.,
 RT "A fibroblast chondroitin sulfate proteoglycan core protein contains
 RL lectin-like and growth factor-like sequences.";
 RL J. Biol. Chem. 262:13120-13125(1987).
 RN [5]
 RP SEQUENCE OF 251-347 FROM N.A.
 RX MEDLINE=93122792; PubMed=1478664;
 RA Iozzo R.V., Naso M.F., Cannizzaro L.A., Wasmuth J.J.,
 RT "McPherson J.D.;
 RA "Mapping of the versican proteoglycan gene (CSPG2) to the long arm of
 RT human chromosome 5 (5q12-5q14).";
 RL Genomics 14:845-851(1992).
 RN [6]
 RP SEQUENCE FROM N.A. (ISOFORM V3).
 RX TISSUE=Brain;
 RL MEDLINE=95181355; PubMed=7876137;
 RA Zako M., Shimomura T., Ujita M., Ito K., Kinata K.,
 RT "Expression of PG-M(V3), an alternatively spliced form of PG-M
 RT without a chondroitin sulfate attachment in region in mouse and human
 RT tissues.";
 RL J. Biol. Chem. 270:3914-3918(1995).
 RN [7]
 RP SEQUENCE OF 3333-3396 FROM N.A. (ISOFORM VINT).
 RX TISSUE=Aortic smooth muscle;
 RL MEDLINE=99327053; PubMed=10397680;
 RA Lemire J.M., Braun K.R., Maurel P., Kaplan E.D., Schwartz S.M.,
 RT "Versican/PG-M isoforms in vascular smooth muscle cells.";
 RL Arterioscler. Thromb. Vasc. Biol. 19:1630-1639(1999).
 RN [8]
 RP PARTIAL SEQUENCE.
 RX TISSUE=Brain;
 RL MEDLINE=89174663; PubMed=2466833;
 RA Perides G., Lane W.S., Andrews D., Dahl D., Bignami A.,
 RT "Isolation and partial characterization of a glial
 RT hyaluronate-binding protein.";
 RL J. Biol. Chem. 264:5981-5987(1989).
 RN [9]
 RP TISSUE-SPECIFICITY OF ISOFORMS.
 RX MEDLINE=96213482; PubMed=8627343;
 RA Paulus W., Baur I., Doures-Zimmermann M.T., Zimmermann D.R.,
 RT "Differential expression of versican isoforms in brain tumors.";
 RL J. Neuropathol. Exp. Neurol. 55:528-533(1996).
 CC -i- FUNCTION: May play a role in intercellular signaling and in
 CC connecting cells with the extracellular matrix. May take part in
 CC the regulation of cell motility, growth and differentiation. Binds
 CC hyaluronic acid.

CC -i- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
 CC -i- ALTERNATIVE PRODUCTS: At least 5 isoforms; V0 (shown here), V1,
 CC V2, V3 and Vint; are produced by alternative splicing.
 CC -i- TISSUE SPECIFICITY: Cerebral white matter. V0 and V1 is expressed
 CC in normal brain, gliomas, medulloblastomas, schwannomas,
 CC neurofibromas, and meningiomas; V2 is restricted to normal brain
 CC and gliomas; V3 is found in all these tissues except
 CC medulloblastomas.
 CC -i- DEVELOPMENTAL STAGE: Disappears after the cartilage development.
 CC -i- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
 CC -i- SIMILARITY: CONTAINS 2 LINK DOMAINS.
 CC -i- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
 CC -i- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -i- SIMILARITY: CONTAINS 1 SUSHI (SCR) DOMAIN.
 CC -i- SIMILARITY: BELONGS TO THE AGGRECAN/VERSICAN PROTEOGLYCAN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 DR EMBL; U16306; AAA65018.1; -
 DR EMBL; X15998; CA34128.1; -
 DR EMBL; S52488; AAB24878.1; -
 DR EMBL; U26555; AAA67565.1; -
 DR EMBL; D32039; BAA06801.1; -
 DR EMBL; J02814; AAA36437.1; -
 DR EMBL; AF084545; AAB48545.1; -
 DR PIR; S06014; S06014.
 DR PIR; A29348; A29348.
 DR PIR; A30358; A30358.
 DR HSSP; P01132; IEGF.
 DR Gene; HGNC:2464; CSPG2.
 DR MIM; 118661; -
 DR InterPro; IPR000152; Asx_hydroxyl.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR000742; EGF_2.
 DR InterPro; IPR001881; EGF_Ca.
 DR InterPro; IPR001438; EGF_II.
 DR InterPro; IPR003599; Ig.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR001304; Lectin_C.
 DR InterPro; IPR000538; Link.
 DR InterPro; IPR000436; Sushi_SCR_CCP.
 DR Pfam; PF00008; EGF_2.
 DR Pfam; PF00047; Ig_1.
 DR Pfam; PF00059; lectin_C_1.
 DR Pfam; PF00084; sushi_1.
 DR Pfam; PF00193; Xlink_2.
 DR PRINTS; PR00010; EGFBL002.
 DR PRODOM; PD000919; Link_2.
 DR SMART; SM00032; CCP_1.
 DR SMART; SM00034; CLECT_1.
 DR SMART; SM00179; EGF_CA_1.
 DR SMART; SM00001; EGF-like_1.
 DR SMART; SM00409; IG_1.
 DR SMART; SM00445; LINK_2.
 DR PROSITE; PS00010; ASX_HYDROXYL_1.
 DR PROSITE; PS00022; EGF_1_2.
 DR PROSITE; PS01186; EGF_2_1.
 DR PROSITE; PS01187; EGF_CA_1.
 DR PROSITE; PS01241; LINK_2.
 DR PROSITE; PS00615; C-TYPE LECTIN_1_1.
 DR PROSITE; PS00441; C-TYPE LECTIN_2_1.
 DR Glycoprotein; Proteoglycan; Lectin; Extracellular matrix; Sushi;
 KW Signal; Repeat; EGF-like domain; Calcium; Immunoglobulin domain;
 KW Hyaluronic acid; Alternative splicing.
 FT SIGNAL 1 20 POTENTIAL.
 FT CHAIN 21 3396 VERSICAN CORE PROTEIN.
 FT DOMAIN 37 137 IG-LIKE V-TYPE DOMAIN.

FT DOMAIN 167 244 LINK 1.
FT DOMAIN 346 346 GAG-ALPHA (GLUCOSAMINOGLYCAN ATTACHMENT
FT DOMAIN 348 1335 DOMAIN).
FT DOMAIN 1336 3089 GAG-BETA.
FT DOMAIN 3089 3125 EGF-LIKE 1.
FT DOMAIN 3125 3163 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 3163 3290 C-TYPE LECTIN.
FT DOMAIN 3290 3353 SUSHI.
FT DISULFID 44 130 BY SIMILARITY.
FT DISULFID 172 243 BY SIMILARITY.
FT DISULFID 196 217 BY SIMILARITY.
FT DISULFID 270 345 BY SIMILARITY.
FT DISULFID 294 315 BY SIMILARITY.
FT DISULFID 3093 3104 BY SIMILARITY.
FT DISULFID 3098 3113 BY SIMILARITY.
FT DISULFID 3115 3124 BY SIMILARITY.
FT DISULFID 3131 3142 BY SIMILARITY.
FT DISULFID 3136 3151 BY SIMILARITY.
FT DISULFID 3153 3162 BY SIMILARITY.
FT DISULFID 3169 3180 BY SIMILARITY.
FT DISULFID 3197 3289 BY SIMILARITY.
FT DISULFID 3265 3281 BY SIMILARITY.
FT DISULFID 3296 3339 BY SIMILARITY.
FT DISULFID 3325 3352 BY SIMILARITY.
FT CARBOHYD 57 57 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 330 330 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 615 615 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 782 782 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 809 809 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1332 1332 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1398 1398 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1442 1442 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1468 1468 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1663 1663 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1898 1898 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2179 2179 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2272 2272 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2280 2280 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2360 2360 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2385 2385 N-LINKED (GLCNAC. .) (POTENTIAL).
Query Match 4.9%; Score 395; DB 1; Length 3396;
Best Local Similarity 19.9%; Pred. No. 1.4e-06;
Matches 431; Conservative 252; Mismatches 725; Indels 756; Gaps 102;
QY 25 VLSGHP-AEASGAGDPADAPAT-----KLPO-----KNQLSVNG 63
DB 786 VLLAHTLSVEAAATVSKSWDEDTTSKPLESTEPSASSKLPPALLTTVMGNKGDIPS 845
QY 64 VARQG--DVHVOENQGBEEVVDVQORESEDRVREKDRVEEMANSTAVEDITKDGQ 121
DB 846 FTEDGADEFTLPDSIQKLEEVTDEDIAHAGKFTIRFQTTSTGIAEKSTLRDST---- 901
QY 122 ETSSEIIQIPASNNV-EMWOPASSQANDVFKVKFV-----GPKFTVKKD 170
DB 902 --TEEKVPPITSTEGQYATMEGSALEVEDVDLSKPVSTVPOFAHTSEVEGLAFVSYSS 959
QY 171 KNE-----KSDTVOLLTVKKDEGEAEASVGAGDH--QEPSVETAVGESAKES----E 218
DB 960 TQPTTVVDSHSTIPUSVPIKTDWGLVLPSPSEDEVLGPSQDILVDIOTRLEATISPE 1019
QY 219 LKQSTKEQGTAKOE-----QSSTEIPLQA-----ESPOAAEERAKDE 256
DB 1020 TMTTKITEGTTQEEFPWKQTAEPVLPALSSANTWTPKEAVTFLDQEGDGSAYTVSEDE 1079
QY 257 ---GEE-----KOEKPTKSPE-----SPSSPVN 277
DB 1080 LLTGSERVPLETTPVGKIDHSVSPGAVTEHKVKTDEVVTLTPRIGPKVSLSPGPEQK 1139
QY 278 SET-----TSSFKFFTHGAGWRKTSFKSKED-DLETAEKKEQEAQKVEEEK 328
DB 1140 YETEGSSTGTGFTSSLPFSTHITQLMBEETTKTSLEIDLGSLFEKPKRATLIEFSTI 1199

QY 329 EKTEPAS-----EQQPAEDTQARLSADYKVELP 359
DB 1200 KVTVPSDIITAFSSVDRLHTTSAPKSSAITTKPPLIDREPGETTDMVIGESTSHVP 1259
QY 360 ---LEQVQ-----DLEASSEKCAPLATE-----VFDEKMEAHQEVVA----- 395
DB 1260 PTTLEDIVAKETDIDREYFTTSSPPATQTPPTPTVEDKEAFQOALSTPQPASTKPH 1319
QY 396 -EVHSTVEKTEREQCGGEAEGGVVVEGTGSLPEKLAEPQEVQEAEPABELM--- 450
DB 1320 PDINVIIIVRENKTRMSD-----LSVIGHPIDSES-KEDEPCSEETDPVDLMAEIL 1372
QY 451 -----KSRECMVGGDHT-----QLTDLSPEEKTLPKHPEG----- 481
DB 1373 PEPPDIIEIDLHYSENEEBECANATDVTTTPSVQYINGKHLVTVPKDPAEAARRG 1432
QY 482 -----IVSEVM-----LSSQRIKVGSS----- 500
DB 1433 QFESVAPQNFSDSESOTHPFVIAKTELSTAVQPNESTETTESLEVTWKPTYPETSEH 1492
QY 501 -----PLKKLFSSSGLKLGKQKGGGGDEEPGEYOHITE----- 540
DB 1493 FSGGEPDVPPTVPFHEEFESGTAKK--GAESVTER-----DTEVGHQAHETEPVSLPPEE 1546
QY 541 -SPESADEQKGESSA-----SSPEPEBETTCLEKPLEAPODGEA-----EGTTS DGE 588
DB 1547 SSGEIALDQESOKIAFARATEVTFGEVEKSTVYTTIPVSSASAYVSEEAVALTIG- 1605
QY 589 KKEGITPM-----ASPKM---VTPKKRV-----RRPESDKBEELKVKSAATL 630
DB 1606 -----NPPWDDLLSTKESVWEATPQVVELSGSSSIPITEGSGEAEDEDTFTMTVDL 1659
QY 631 SSTDSF-----VSEMODEVKTGVEEQKPEPKRV-----DTSVSEALI 670
DB 1660 SORNTTDTLITLDSRIITESFEVPATTIYVSEQPSAKVPTKVFSETDTS-EMISST 1718
QY 671 CVGSSKKRARKASSDDEGGPRTLGLDGHRAEASKEAGTDAVPASTQEOQOQ---AQG 727
DB 1719 TVEEKKKK-----EEGTTGT-----ASTFEVYSSQORSQDLILPPE 1755
QY 728 SSSPEPAGSPGEGVSTWESFKRLVTPRKKSKELEK---AEDSSVOLSTEIPSRE 784
DB 1756 LESPNVATSSDSG---TRKFSMLTTPQSEREMTDSPTVFTTNTLENLGAQ---TTE 1808
QY 785 ESWVSIKKFIPGRKKRADGKOQATVEDSGRVEINED-----DPNVPAVPLS--- 833
DB 1809 HSSI-----HQPVGEGTLTPRSPASVFMEOGSGEAAADPETTTVSSFSLVN 1856
QY 834 EYNAVBREK-----MEAGNTE-----LPQLLAGVYVSEELSK- 866
DB 1857 EY-AIOAEKEVAGTGLSPHVTFTSTPTGLVLTVMDRVVAENITQTSREIVISERLGE 1915
QY 867 -----TLVHTVVS-----VAVIDGT-RAVTSVEERSPSWIS 895
DB 1916 NYGAEIRGFTGFPLEBEDFGDFREYSTVSHPIAKEETVMMEGSGDAFRDQTSPTSTVP 1975
QY 896 ASVTPELHTAGEAMP-----PVEVTE-----KDIIAETPVLQTLPKGDA 939
DB 1976 TSV--HISHISDEGSPSTWVSTSAFFBEEFTSSAEGSGEQLVTVSSVW-PVLPASVQK 2032
QY 940 HDDMVTSEVD-FTSEAVTATE---TSEALRTEEV---TEASGAETTDMVSAVSOLDTSPD 993
DB 2033 FSGTASSIIDEGIGEVGTVNEIDRRSTILPTAEVGTAKPAVEEKVKGSTVS--TNFPQ 2090
QY 994 TTE-----EATPV-QEVESGVLDEEBEERQTOAILQAVADKVKESQVPATQTVORT 1044
DB 2091 TIIEPAKLWSRQEVNVRQIESET--TSEQIQEEKSPESPONSATEQTIIPDSQIFFTET 2148
QY 1045 GSK-----ALEK----- 1051
DB 2149 ELKTTDYSVLTKTKTYSDDKEMKEEDTSLVNMSTPDPDANGLESYTTLPEATEKSHFFLA 2208

```

QY 1052 -----VEEVEDSEVLASEKEDVWPKG---PVQAGAEHLAQSEGTQAT----- 1094
Db 2209 TALVTESIPAEHVVDSPKKESTKH-FPKGMPTIOESDTELLFSGSGSEVLP 2267
QY 1095 PESLEVPV-----TADVDHATCOVKLOOL-MEOAVAPESSET----- 1133
Db 2268 TESVAFTEVEQJNNTLYPHYTSQVESTSKIDPFNMENAVAKEVGLVSGTIFEGSGSV 2327
QY 1134 -----LDTSETNGST--PLADSDPTADGTQODETIDSQSKATAAARQSOVTEEE-- 1180
Db 2328 TSTTLIELISDYGAGPVTAPLPFS-TDIGHQONQVRWAEBIQIS--RPGITTEDQSNK 2384
QY 1181 -AATQKEPSTLPNNVPAQEHGEPGCDVLEPPQOELTAAPPLATETVQGE--GEV 1237
Db 2385 NSTAEINETTISTDFLAR-AVG-----PEMAKEFVTSAPKBSDLVEYSGSGSGEV 2436
QY 1238 DMLDEKV-----KEQEVFVHSG-----PNSOKADVTVYD-----SEWVAGVQCE 1279
Db 2437 DIVDSFHTSATTOATROESSTTFVSDGSLKHPVPSAKAVADGFPVSVWMLPLHSEON 2496
QY 1280 KESTEVQSL--SLEEGEMETDVEKEKRETKPQVSEGEQETAAPHEGTYGKPVLLDM 1337
Db 2497 KSSPPTSTLSTMTVGERTDGSPQDR-----PREFEDSTLKPRKKPTENIITDLK 2549
QY 1338 PSSERKALGSLGSGPSLPD--QDRACTEV-----QVQSLDTTV----- 1375
Db 2550 EDKDLILITTESITLLEIPELTSKDXTLIDIDHTKPVYEDILGMQTDITEVPSPHDSN 2609
QY 1376 -----TQTAFAVEKIVTVVISERGE-SPECVGAHLPAKKSAT-----GGH 1417
Db 2610 DESNDSTQVQIYEAVNLSTRETFEGSADVLASTQATHDESNITYDRQGLDHMGPH 2669
QY 1418 WTLQHAEDTVPEGPESQAE-----SIPITVPAPESTLHPDQG----- 1456
Db 2670 FT-----TGIF-APSTETELVDLPLPATSLPI--PRKATVYIPELGGKAEKALDME 2720
QY 1457 -----EI--SASQREKSEF--EDKPDAGDADGKESTALEKVL----- 1490
Db 2721 ESSTLSDGQALDOSERILPTLQCFERTOEYEDKKHAGSPQEPFSSGAEELVDHTPYL 2780
QY 1491 -----KAPPELLLEKSNKILVNIQTAVDQFAR-----TERAPETHAVDSQ----- 1533
Db 2781 SIATTHLMDOSVTEVPDVEGNSNPYYTDTTLAVSTFAKLSQTPSSPLTIISGSBASGH 2840
QY 1534 TQVP 1537
Db 2841 TEIP 2844

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ch/announce/
CC or send an email to license@ebi-sib.ch).
CC -----
DR EMBL; M34384; AAA41685.1; -.
DR PIR; A34736; A34736.
DR InterPro; IPR001664; IP.
DR Pfam; PF00008; filament; 1.
DR PROSITE; PS00226; IF; 1.
DR K1 Intermediate filament; Coiled coil; Neutrone.
FT DOMAIN 1
FT DOMAIN 8 314 ROD.
FT DOMAIN 315 1805 TAIL.
FT DOMAIN 8 43 COIL 1A.
FT DOMAIN 44 55 LINKER 1.
FT DOMAIN 56 151 COIL 1B.
FT DOMAIN 152 174 LINKER 12.
FT DOMAIN 175 193 COIL 2A.
FT DOMAIN 194 196 LINKER 2.
FT DOMAIN 197 314 COIL 2B.
SQ SEQUENCE 1805 AA; 198744 MW; B40BE14717E0998D CRC64;

Query Match 4.8%; Score 388; DB 1; Length 1805;
Best Local Similarity 21.3%; Pred. No. 1,38-06;
Matches 319; Conservative 233; Mismatches 543; Indels 402; Gaps 75;

QY 13 EOPAGSDF--PSELV-----LSGHGPAEASGAADPADDPATKLPQKNGQLSSVNGV 64
Db 440 BEPGKQOQHPPDDLTSATNLNPHHPTLEAKD--GESSESRSVISIFQEDSGQIMEL--V 495
QY 65 ABQGDVHVOENQEGQ--EEEVVDQVQRESEDDVREKDRVEMAN--STAVEDIRKD 119
Db 496 EKEADIEKVENNSAQKTOESGLDTEFTQ--DSQGLQKTLKALGEBPLMSLKTIONYETA 554
QY 120 GQE-----ETSEIIEQIP--ASENNVEWVQPAESQANDVCFKVPYFVG 163
Db 555 GKENCNSSTBGLTGLBEPEKEKQJPLKSLBKVNE-----SEKTLNGVPVLSSELLGK 608
QY 164 KFTVKKDKNEKSDTVQILITVKKDSGEAGASVAGADHOEPPVET-----AVGESA 213
Db 609 EDTRTEDQELMSPKG--TLKRFSSLGKESQEVVAPRSKGNLSEWTAKESQHPGLFPG 665
QY 214 SKESLXKSTKQEBGL--KQBSSTEIPLOAESDQ--AAEEKAKDGE-----EKQE 262
Db 666 AEDQMLERLVEKEDQSPRSPPEEDQACRPLQKNOEPLGYEAA--BQQLIERLIEKES 723
QY 263 KEPTSPSPSSPVNSETTSFKEFTTGAWGWRKTKTSFKSKEDDLTAERKKEQEA 322
Db 724 QESLRSPEEB--DQEAQSLQK-----ENQRPPLGTEAEADQMLRLIEKESQESL 772
QY 323 VDEBEKTEPASEQOE-----PAEDTDQARISA-----DYEKVLEPLR-- 361
Db 773 SPEENORICKPLERENQKSLRYLEENQETFPVLEBRNORPLRSLVEBEQIVAPLEKV 832
QY 362 --DQVGDLEASSEKCAPLATEVPFEKMEAHQVAVHVSIVKTEBEGGGGAGBAGV 419
Db 833 SODSLGSL--AEENVQPLR--YLEEDDCJNKSLE-----DKTHKSLGSLERBNGDS 880
QY 420 VV-----EGTGESL--PPEK-----LAPEQVPOEAPABELMKSRMCVSGGDHTOL 465
Db 881 IITPQESQVSLRPPBEDQRIYVHLKESQEPFRSSBEEQVW--BRLSEGNHBL 937
QY 466 TDLSPB--EKLTPKHDPGIVSEVEMLSQGRIRIVQSGPLKLFSSGGLKKLGGKQK 521
Db 938 SSVKEDQWESQLKESQDSGLSD--ESQETF--GPLEK--BNASIRSLAQDQ-- 989
QY 522 KKGGGDEBPGYQIHHSPESADQKESASAPSEEBETTCLEKPLRPOD--GEAF 580

```



```
Db 990 -----ESQKLEQETQOTLRVGNQMAVSP-----EKVDPELPKPLGNDQ 1030
Qy 581 EGTTSDEKREKRGITWASPKWVTPK-----RVRRPSESDEKBELEKVS-----ATLSST 633
Db 1031 EIARSLGKNEQSLVSLKE-KGIETVKSLETEIIEPLET-AEEDLERKSIDTQEPWLST 1088
Qy 634 D-----STVSEMQDE-----VKTVGEEQRP-----EPPKRVDTSVSWEALICVSGSKKARKA 682
Db 1089 EVARETVPEDEPPGSLGSDVNRRETLTSLKESQELSLGKWNVETRVEDSQOCLQVE 1148
Qy 683 SSDDGGGPRTL-----GDSHRAEASDKKAGTDAVPAST-----719
Db 1149 EGIQEQHQESLEVKQELPSSGNQORWEDVVGKAVGQAPLATTGVTGDKAEHLRG 1208
Qy 720 ---QEODQAQ-----GSSPEPAGSPSEG-----EGVSTWESFKRLVTPRK 758
Db 1209 QGGEERAAAGELLQDITVGEAWSLGSEPKQRPVPAEALDNLEGA-----LEVPAQ 1261
Qy 759 SKSKLEEKAD--SSVEOLSTEIEPSREESWVSIKKFIPGRKRKRGADKQEQATVEDSGP 816
Db 1262 SDEVTERDEDRAGQEQDSIEVTLGLEAARTGLEL-----EQQVVGLED--P 1307
Qy 817 VEINEDPNPVPVPSVNAVREKMEACQNTLPPOLLGAVTVS-EELSKTLVHTVSVVA 875
Db 1308 RHPAREEAIPPS---LGEESVKAKIAQLEGPGCKEKGALDGLSLGILPKTSEALECQ 1364
Qy 876 VIDGTRAVTSVERSPSWISASVTEPLEHTAGBAMPVPEVETKDIIAET-----926
Db 1365 GHEESEMGEWEEAAS-----LETSDHEGSDAPQPRPPETEDEGAQAALTAPGPKLL 1418
Qy 927 -----PVLITQTLPEGKADHDMVTSE-----VDFTSEATATETSEALTEEVTEA 972
Db 1419 EPCSPILIT-----DAHELOPQAEIGIQEAGWQPEAGSEALERVENEPFGLGPIEG 1471
Qy 973 -----SGAETTDWMSAVSOLTPSPDTEATP-----VQEVESGLVDTTEEBEERTQAILQ 1023
Db 1472 LQWEGREE-----SEADLGETLPDSTPLGLVLRSPAPKWDLAGQR-----LS 1518
Qy 1024 AVADKKEE--SOVPAQTQVTRGSKALEKVERVEEDSEVLASEK-----EKDVMKGP 1075
Db 1519 PQGDACKEDWGPVPAQA-----GLSPPEEEEOGHGSDLSSEEFEDLGTASLLPGVP 1573
Qy 1076 VQAGAEHLAQ-----GSETQATPESL--EVPEVTADVHVAT-----1112
Db 1574 KEY--ADHVGQVPPVLQPCWQDGESDGFADBEESGEEGEDADEGAESGAQWWSG 1631
Qy 1113 -----COVIKLOLMQEAQVAPESSETLTQSETNGST-----PLASDSDTADGTQODETID 1161
Db 1632 ASGGGCKVQDIAQRGDPVQESVGVSLGWDGLRGAANVPALEMVQSQAEPGSSES-- 1689
Qy 1162 SQDSKATAAVROSQVTEEEAATAQKEEPSPLPNNVPAQEHBGEPGRDVLPTQBELTAA 1221
Db 1690 ---ESASLEEGEQVTDH--LDAPQEVTSVMPGVGDAFDIGGQSPNLD-----1732
Qy 1222 AVPLAKTEVG---QSGEVNLDGCKVKERQEVF-----VHSGPN 1258
Db 1733 SEQVGMKENGLEQAEQGV-VLDGD---EQELLQOQEVGALKPLVLPASPVHLGPS 1785

RESULT 14
MAP2_RAT
ID MAP2_RAT STANDARD; PRT; 1861 AA.
AC P15146;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Microtubule-associated protein 2 (MAP 2) (MAP2B) [Contains: MAP2C].
GN MAP2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
```

```
RP SEQUENCE OF 1-1694 AND 1726-1861 FROM N.A.
RC STRAIN=Wistar; TISSUE=Brain;
RX MEDLINE=90251471; PubMed=2339070;
RA Kinder S., Schwanke B., Schulz B., Garner C.C.;
RT "Complete cDNA sequence encoding rat high and low molecular weight
RT MAP2.";
RL Nucleic Acids Res. 18:2822-2822(1990).
RN [2]
RP SEQUENCE OF 1-1694 AND 1726-1861 FROM N.A.
RC STRAIN=Wistar; TISSUE=Brain;
RX MEDLINE=91060576; PubMed=2174050;
RA Kinder S., Schulz B., Goedert M., Garner C.C.;
RT "Molecular structure of microtubule-associated protein 2b and 2c from
RT rat brain.";
RL J. Biol. Chem. 265:19679-19684(1990).
RN [3]
RP SEQUENCE OF 1-151; 1515-1694 AND 1726-1861 FROM N.A.
RX MEDLINE=90221819; PubMed=2326166;
RA Doll T., Papandriopoulou A., Matus A.;
RT "Nucleotide and amino acid sequences of embryonic rat MAP2c.";
RL Nucleic Acids Res. 18:361-361(1990).
RN [4]
RP DISCUSSION OF SEQUENCE.
RX MEDLINE=89365159; PubMed=2770869;
RA Papandriopoulou A., Doll T., Tucker R.P., Garner C.C., Matus A.;
RT "Embryonic MAP2 lacks the cross-linking sidearm sequences and
RT dendritic targeting signal of adult MAP2.";
RL Nature 340:650-652(1989).
RN [5]
RP SEQUENCE OF 1695-1725 FROM N.A.
RX MEDLINE=94110302; PubMed=8282767;
RA Doll T., Meichner M., Riederer B.M., Honegger P., Matus A.;
RT "An isoform of microtubule-associated protein 2 (MAP2) containing
RT four repeats of the tubulin-binding motif.";
RL J. Cell Sci. 106:633-640(1993).
CC -!- FUNCTION: THE EXACT FUNCTION OF MAP2 IS UNKNOWN BUT MAPS MAY
CC STABILIZE THE MICROTUBULES AGAINST DEPOLYMERIZATION. THEY ALSO
CC SEEM TO HAVE A STIFFENING EFFECT ON MICROTUBULES.
CC -!- ALTERNATIVE PRODUCTS: VARIOUS FORMS OF MAP2 ARE PRODUCED BY
CC ALTERNATIVE SPLICING OF THE SAME GENE. MAP2C, THE LOW MOLECULAR
CC FORM OF MAP2, LACKS THE CENTRAL DOMAIN OF MAP2A/B.
CC -!- DEVELOPMENTAL STAGE: MAP2C IS EXPRESSED DURING EMBRYONIC BRAIN
CC DEVELOPMENT AND UNTIL POSTNATAL DAY 10. MAP2B IS EXPRESSED
CC THROUGHOUT BRAIN DEVELOPMENT.
CC -!- SIMILARITY: CONTAINS 3 OR 4 TAU/MAP REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X51842; CAA36135.1; -.
CC EMBL; X17682; CAA35667.1; -.
CC EMBL; X71487; CAA50588.1; -.
CC PIR; S07887; S07887.
CC PIR; S10003; S10003.
CC PIR; A37981; A37981.
CC InterPro; IPR001084; Tubulin_Tau.
CC Pfam; PF00418; tubulin-binding; 4.
CC PROSITE; PS00229; TAU MAP; 3.
CC
CC Microtubules; Repeat; Alternative splicing; Calmodulin-binding.
CC DOMAIN 1454 1474 CALMODULIN-BINDING (POTENTIAL).
CC REPEAT 1664 1694 TAU/MAP MOTIF.
CC REPEAT 1695 1725 TAU/MAP MOTIF.
CC REPEAT 1726 1756 TAU/MAP MOTIF.
CC REPEAT 1757 1788 TAU/MAP MOTIF.
CC FT VARSPPLIC 152 1514 MISSING (IN ISOFORM MAP2C).
CC FT VARSPPLIC 1695 1725 MISSING (IN ISOFORM WITH 3 TAU/MAP
CC REPEATS).
CC SEQUENCE 1861 AA; 202409 MW; 42DCF116D21EF54E CRC64;
```

Query Match 4.8%; Score 388; DB 1; Length 1861;
 Best Local Similarity 19.8%; Pred. No. 1.3e-06;
 Matches 306; Conservative 220; Mismatches 590; Indels 432; Gaps 66;

QY 7 TQORPEQAGSDTPESEVL-----SGHPPAABASGAAGDPADDPATKLQKQKQSSVNG 63
 DB 335 TQSGOPFAVPFQSDDKMSLQDTSGATSKES-----KDEBPQDKADKADVVSZA 388
 QY 64 VAEQGVHVE-----ENQEQEEVEVDVQORESEVDREKRDVEEMANSTAVEDIT 117
 DB 389 TTVLGDVHSPAVBGVGENISGEKGTTOQEKKETSPVQOE-----PLTETEPQT 440
 QY 118 KQGEETSEIIQOIPASENNVEMVOPASQANDVGFKKVFGKFTVKDKDKNEKSDT 177
 DB 441 K--LEETSKV-----SIEETVAKKEESL-----TKDKDQAV 470
 QY 178 VQLLT---VKQDEGEGEASVGA--GDHQBPSVETAVGSASKESELKOSTEKQEGTLKQ 232
 DB 471 IOTSTQSFSGEDQKQOEOTIEALKQDSFPISEQAVTDMAATKLTLEKVTSEPEAVSEK 530
 QY 233 EOSTEIPLOAESDQAAEEAKDEGEKQEKPTKSP--ESPSSPVNSEBTSSFKKFTTH 290
 DB 531 RE-----IQGFEEIDIAKSKLEGAGSATVAEVEMPFYEDKSGMSKYFETSAKDEDVTR 584
 QY 291 G---WAGMRKTSFKKSKEDDLTAEKKEQAEKVDEBEKTEBPASEBOEPAETDQA 347
 DB 585 STGLGSDYVELSDSRGNAOESLDTVSPKNOQ-----DEKILLAKASQPSPPAHAGYS 637
 QY 348 RLSDAV---EKVELPLEDDQVGLASSEKCAPLATEVEDEKMAEQVAAEHVSTVK 404
 DB 638 TLAOSTSHPELPEE-----PSSPQERMTFTDPKVYGEKKDLHKKNDLTLT--- 687
 QY 405 TEEEOGGGEAGGAVVE-----GTGESLPPEKLAEPQEVPOE 442
 DB 688 --RSLGLGGRS--AIEGRSMSINLPMGCLDSIALGFNGRGHDLSP----- 729
 QY 443 ABPAEELMKSREMCVSGDHTQTLTDSPEKTLPKHPEGIVSEVEMLSOERIKVQGSPL 502
 DB 730 --LASDILINTSGMDEGD--DYLPTTPPAVEKIPCP--IESKEBEDTEQAKVTGQOT 783
 QY 503 KTLFSSGKLKLSGKKQKGGGGGGBEPGEVYHHTESPESADQEGSSASPEEPE 561
 DB 784 TQVETSS-----ESPFPKEXYKNGTVMAPLPEH 813
 QY 562 -----ETTCLEKGPLLEAPQDEAEGETTSDDEKREGITTPWASFKKAV 604
 DB 814 LDLAGRSKLASASADAEVARRKVSSEAV---VAESST-----GLPPVADDSQPV 861
 QY 605 TPKKRVRRPESDKEELEKVSATLSS--TDSTVSEMDEBKTVG---EOKPEEPPK 658
 DB 862 KP-----DSQLDMDGVCVFNKTYVPLPSPVQDSENLSGSGSFYEGTDDKVR 909
 QY 659 RVDTSVSMALICVG--SSKKRAKAKASSDDEGPRTLGGSHRAEASDKDEAG--TDVAV 715
 DB 910 DLTATDLS---LLEVKLAAAGRVADDEFTAEKEASPPSSADKSGISRFPPDQKRAKNDLDTV 966
 QY 716 PASTQO--DOAQSSSPPEAGSPSE--GEGVSTWE--SFKRLVTPPKKSKLEKAED- 769
 DB 967 LEKSEHNVSKELAKSEBEVGDVLEFLGLGV--TYEQTSKAKELT-----TKETAPERAAKNG 1021
 QY 770 -SSVEQLSTEIEPSREESVWSIKK1PGRKKRADKQEOATVEDSGPVEINEDD----- 823
 DB 1022 LSSVPEVA-EVETT-----TKADQGLDVAKKDDQSDPLDIDKVSDEGQMA 1064
 QY 824 --PNVAVAVPLS-EYNAVEREKMAQNTLPLQLG-----AVYVSELSK- 866
 DB 1065 SGMSVDAGKTIELKFEVDQOLTLSSAPQETDSFMGISSHVADGAKVSETEKENVAKR 1124
 QY 867 TLVHTVSV-----AVIDGTRAVTSVEERSPSM-----ISAGVTEPLEH 904
 DB 1125 DLVHQAVDKESYESSGHESLTMSLAKPDEKKEITSPETSLIQDBVALKLSVEIPLCP 1184

QY 905 TAGEMPPVEEVEYTEKDI-----IAETPVLTQTLPEGKXAHDDMTSEVDTSEATATET 960
 DB 1185 PVSEADSSIDEAEVQMFQIQLPKEESTETPDIP-----ALPSTDVQPOPEAVVSEP 1236
 QY 961 SEALTEEVEENSGAEFTTDMWASVSLTD--SPDTTEATPVQEVESGVLDTEEBEERQT 1018
 DB 1237 AEVREEEIEIEAGEYDKLRLSRDITQITDILVPSRREF-----VETCPBEH 1284
 QY 1019 QAILQAVADKVKESQVPATQTVORTGSKALEVE-----EVEDESVLASEKEDVMPK 1073
 DB 1285 KGVSESVV--TLEDPFITVQTTDEGELGSHSVFAFAPVQPEEERRRPPHDELEVL-- 1340
 QY 1074 GPVQAGAEHLAQSGETQATPESLEVEPVATADVHVATQVYIKLQQLMEQAVAPESSET 1133
 DB 1341 -MAAEQAQAP-KDGSPPDAIPATPEKEVP-----PSEYKTEY 1374
 QY 1134 LTD--SETNGSTPLADSDPTADQTOQDETIDSQDSKATAVAQSQVTEBEAATAQKEPST 1191
 DB 1375 YDYKDEFTTIDSIINDASL-----WVDYQDD-----DRSLTQLEITIPKEERAE 1420
 QY 1192 LFNVPQAQEEHGE--PGRD-VLEPTQOELTAAAVPLAKTEYQGEVDMLDGEVK 1246
 DB 1421 KEARRPSLEKHKRKEKPFYTGGRGRISTPERRREYAKKEPSTVSRDEVRKKAV--YKKAELA 1478
 QY 1247 BEQEVFVHS-----GPNQKADAVTYDSEV 1271
 DB 1479 KESEVQAHSPSRKLLKPAIKYTRPHTLSYVKRTTATSGESAQAPSAFKQAK--DKVT 1535
 QY 1272 MGVACCOCESTEVQSLT-----BEG--EMETDVEKEKRETKPEQVSEEGOE 1318
 DB 1536 DGITSPKRSLLPRPSSILPRRGVSGDRRENSLWSSISSARTTRSEPIRAGSG 1595
 QY 1319 TAPEHEGTYGKPVLTLDMPSSERKALGSLG--GSPSLPQDKACGT 1364
 DB 1596 TSTPT---TPGSTAITPGTSPSSRTPTGPTPSPYPRTPGHPKSGIL 1640

RESULT 15
 NFH MOUSE STANDARD; PRT, 1087 AA.
 AC P19246; O61959;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Neurofilament triplet H protein (200 kDa neurofilament protein)
 DE (Neurofilament heavy polypeptide) (NF-H).
 GN NFH OR NFH.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 NX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89121513; PubMed=3220257;
 RA Julien J.-P., Cote F., Beaudet L., Sidky M., Flavel D., Grosfeld F.,
 RA Mushynski W.,
 RT "Sequence and structure of the mouse gene coding for the largest
 RT neurofilament subunit.",
 RL Gene 68:307-314 (1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89089138; PubMed=3145094;
 RA Sheidman P.S., Garden M.J., Lees J.F., Iazzarini R.A.,
 RA "The structure of the largest murine neurofilament protein (NF-H) as
 RT revealed by cDNA and genomic sequences.",
 RL Brain Res. 464:217-231 (1988).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Swiss Webster; TISSUE=Brain;
 RA Garden M.J.;
 RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M,
 AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.

CC NF-H HAS AN IMPORTANT FUNCTION IN MATURE AXONS THAT IS NOT
 CC SUBSERVED BY THE TWO SMALLER NF PROTEINS.
 CC -!- PTM: THERE ARE A NUMBER OF REPEATS OF THE TRIPEPTIDE K-S-P, NFH IS
 CC PHOSPHORYLATED ON A NUMBER OF THE SERINES IN THIS MOTIF. IT IS
 CC THOUGHT THAT PHOSPHORYLATION OF NFH RESULTS IN THE FORMATION OF
 CC INTERFILAMENT CROSS BRIDGES THAT ARE IMPORTANT IN THE MAINTENANCE
 CC OF AXONAL CALIBER.
 CC -!- PTM: PHOSPHORYLATION SEEMS TO PLAY A MAJOR ROLE IN THE FUNCTIONING
 CC OF THE LARGER NEUROFILAMENT POLYPEPTIDES (NF-M AND NF-H), THE
 CC LEVELS OF PHOSPHORYLATION BEING ALTERED DEVELOPMENTALLY AND
 CC COINCIDENT WITH A CHANGE IN THE NEUROFILAMENT FUNCTION.
 CC -!- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
 CC -!- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN IN POSITIONS 534
 CC TO 716 AND IS SHORTER DUE TO FRAMESHIFTS.
 CC
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M24496; AAA39813.1; --
 CC EMBL; M23349; AAA39813.1; JOINED.
 CC EMBL; M24494; AAA39813.1; JOINED.
 CC EMBL; M24495; AAA39813.1; JOINED.
 CC EMBL; M35131; AAA39809.1; JOINED.
 CC EMBL; M35131; AAA39809.1; ALT_FRAME.
 CC EMBL; M31012; CAA83229.1; --
 CC PIR; JTO3368; QFM5H.
 CC PIR; A43778; A43778.
 CC MGD; MGI:97309; Nfh.
 CC InterPro; IPR001664; IF.
 CC Pfam; PF00038; filament; 1.
 CC PROSITE; PS00226; IF; 1.
 CC Intermediate filament; Coiled coil; Neurone; Phosphorylation;
 KW Repeat.
 FT DOMAIN 1 97
 FT DOMAIN 98 408
 FT ROD.
 FT ROD.
 FT TAIL.
 FT TAIL.
 FT GLU-RICH (ACIDIC).
 FT 50 X 6 AA TANDEM REPEATS OF K-S-P-A-E-A.
 FT GLU/LYS-RICH.
 FT COIL 1A.
 FT LINKER 1.
 FT COIL 1B.
 FT LINKER 12.
 FT COIL 2A.
 FT LINKER 2.
 FT COIL 2B.
 FT K -> QA (IN REF. 2 AND 3).
 FT A -> AR (IN REF. 2 AND 3).
 FT S -> T (IN REF. 2 AND 3).
 FT L -> G (IN REF. 2 AND 3).
 FT P -> PREAKSP (IN REF. 3).
 FT MISSING (IN REF. 3).
 FT V -> A (IN REF. 3).
 FT G -> M (IN REF. 2 AND 3).
 FT T -> N (IN REF. 2 AND 3).
 FT T -> N (IN REF. 2 AND 3).
 FT 116612 MW; 57BAC76A38ED1CB9 CRC64;
 SQ SEQUENCE 1087 AA; 116612 MW; 57BAC76A38ED1CB9 CRC64;
 Query Match 4.8%; Score 385.5; DB 1; Length 1087;
 Best Local Similarity 23.8%; Pred. No. 9.3e-07;
 Matches 219; Conservative 139; Mismatches 374; Indels 189; Gaps 43;
 QY 26 LSGHGA-AEASGAGDPADAPATKLPQNGQLSVNGVAEGDVHVOE-----ENQEQ 80
 DB 240 IQCGGAQAQAQAARDALKCDVTSALRIRAQL---EGHAVQSSLSQSEWFRVRLRLS 296
 QY 81 EEEVDDVQRESDVREKDRVEEMAANSSTAVEDI--TKDQGE-ETSEIIIEQIPASENN 137
 DB 297 EAAKVTNDAMRQAQEITEYRR--QLQARTTEALKSTKESLERQSRSELEDRHQADIAS 354

QY 138 VEMVQPAESQANDVGFKKVFKVGFVKFTVKKDKKNEKSDTVQLLTVK-----KDE 187
 DB 355 YQDAIQQLDSELRN-----KWEMAAQLREYQD---LLNVKMALDIEIAAYRKL 400
 QY 188 GGAASVAGAGDH-----QEPSVETAVGESAKSELKOSTEQ-----EGTLKQEQ 234
 DB 401 LGECECRIGFGSPSLTEGLPKIFSIITHIKVSEMIKVVKESEKIVVEGGQTEER 460
 QY 235 SETEIPQAESD-QAAEEBAKDEGEKQKE---PTKSPESPSPVNSNETTSFKFFTH 290
 DB 461 VTGVTVEEDKQAQGEAGEAGEEKEEELAAATSPAAEAAASPEKETSRVKE--- 516
 QY 291 GWAGWRKTSFPKSKEDDLLEAKKEQPAEKVDE-----EKEKTEPASEQEPADTD 345
 DB 517 -----EAKSPGEAKSPGEAKSPGEAKSPGEAKSPGEAKSPGEAKSPGEAKSP 568
 QY 346 QARLSADYKVELPLEDQVDLEASSEKCAPLATEVDFDEKMEAHQEVVAHVSVTEK 405
 DB 569 PAEAKSPAEKSPATVKSPEAKSPAEKSPAEKSPAEKSPAEKSPAEKSPAEKSPA 622
 QY 406 EEEQGGGGAEGGVVVGTEGSLPPEKLAEPQEVQEAEE---PAEELMKSRMCVSGGDH 462
 DB 623 PAEAKSPAEKSPATVKSPEAKSPAEKSPAEKSPAEKSPAEKSPAEKSPAEKSPA 680
 QY 463 TQLTDLSPPEKTLPKHPEGIVSEVEMLSQE-----RIKVGQ---SPLKKLFSSSGLKKL 514
 DB 681 K-----SPAEE---PKSPAEAKSPAEKSPAEKSPAEKSPAEKSPAEKSPAEKSP 731
 QY 515 SCKKQKKGKGGGD-----EEPGCYOH-IHTESPESADE--OKGESASSPEPEETT 564
 DB 732 AAVKSPGEAKSPGEAKSPAEKSPAEKSPAEKSPAEKSPAEKSPAEKSPAEKSPA 791
 QY 565 CLEKGPLEAPQDGEAEGTTS---DGEKKREGITPW-----ASFKGMVTPK 607
 DB 792 SPVKEDIKPPAEKSPAEKSPAEKSPAEKSPAEKSPAEKSPAEKSPAEKSPAEKSPA 851
 QY 608 KVVRRPS-ESDKEELEKVKVATSLSTSTVSEMDEVKT-VGEBOQKPEEPKRVDTSVS 665
 DB 852 EQVKSPAKEAKSPKEEAK-----TSEKVAPKKEEVKSPVKEEVKAKEPPKVBEE-- 903
 QY 666 WEALICVSSKKRANKASSDDDEGGPRTLGGDSHRAEASDKDEAGTDVAPASTQDQA 725
 DB 904 -----KTLPTPKTEAKSSKD-EAPKEAPKPKVBEKKE- 935
 QY 726 QGSSSPGAGSPGEGVSTWESFKRLV-----TPRK---KSKSLEEAEDSSVEQLSTE 778
 DB 936 --TPTEKPKDSTAEAKKEAGEKKAVASEETPAKLGKVEAKPKETTKTTAEADTK 993
 QY 779 I-EPSR-BESWVSIIKFIPIGRKKKADQKQEQATVEDSGPVE-----INEDDPNPVAV 829
 DB 994 AKEPSKPTETEKPKKEEMPAAPAEKK-DTKEEK-TTESRKPKEKPKWEAKVKEDDKSLS-- 1049
 QY 830 VPLSEYNAVEREKMEAQNGTE 850
 DB 1050 ---KEPSKPKTEAKSSSTD 1067

Search completed: December 13, 2002, 00:13:37
 Job time : 85 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 13, 2002, 00:00:30 ; Search time 94 Seconds

(without alignments)
3498.417 Million cell updates/sec

Title: US-09-902-432-4

Perfect score: 8073

Sequence: 1 MCAGSSTQRSPEQAGSDT.....AWAQRKCLPRLQKAPVSK 1596

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8073	100.0	1596	11 Q62766	Q62766 rattus norv
2	8030	99.5	1596	11 Q921F7	Q921F7 rattus norv
3	6512.5	80.7	1684	11 Q9WTQ5	Q9WTQ5 mus musculus
4	6035.5	74.8	1579	11 Q9MP1	Q9MP1 mus musculus
5	1805	22.4	2471	13 Q9DDN8	Q9DDN8 xenopus lae
6	642	8.0	2768	5 Q9VC00	Q9VC00 drosophila
7	635	7.9	5458	5 Q9U459	Q9U459 plasmodium
8	598	7.3	17352	5 Q95YM2	Q95YM2 procambarius
9	585.5	7.3	1829	5 Q22248	Q22248 caenorhabdi
10	574	7.1	5327	5 Q76891	Q76891 drosophila
11	564.5	7.0	2910	10 Q9FND5	Q9FND5 arabidopsis
12	558	6.9	16215	5 Q9NFS3	Q9NFS3 drosophila
13	557.5	6.9	6815	5 Q917U4	Q917U4 drosophila
14	517	6.4	2081	10 Q91H98	Q91H98 arabidopsis
15	501.5	6.2	3484	5 P91257	P91257 caenorhabdi
16	493.5	6.1	2083	5 Q9N435	Q9N435 caenorhabdi

17	489.5	6.1	1852	3 Q9C2H4	Q9C2H4 neurospora
18	489	6.1	6632	5 Q01761	Q01761 caenorhabdi
19	489	6.1	6994	5 Q17343	Q17343 caenorhabdi
20	488	6.0	5170	5 Q17490	Q17490 caenorhabdi
21	485	6.0	2484	5 Q9U347	Q9U347 caenorhabdi
22	485	6.0	2607	5 Q23187	Q23187 caenorhabdi
23	485	6.0	3263	5 Q917U3	Q917U3 drosophila
24	485	6.0	6632	5 Q17362	Q17362 caenorhabdi
25	472	5.8	1514	5 Q85Y55	Q85Y55 drosophila
26	466.5	5.8	1871	10 Q9SRD5	Q9SRD5 arabidopsis
27	463	5.7	1661	5 Q06166	Q06166 plasmodium
28	462	5.7	1786	5 Q9U0P0	Q9U0P0 plasmodium
29	456.5	5.7	1804	11 Q9Z1Q1	Q9Z1Q1 mesocricetu
30	452	5.6	3239	5 Q8T1Q2	Q8T1Q2 bombyx mori
31	442	5.5	1326	16 Q9L2C3	Q9L2C3 streptomyce
32	440.5	5.5	1812	5 Q9VZD9	Q9VZD9 drosophila
33	438	5.4	1128	5 Q26947	Q26947 trypanosoma
34	438	5.4	1616	5 Q9VSJ0	Q9VSJ0 drosophila
35	434	5.4	1510	5 Q25920	Q25920 plasmodium
36	431	5.3	3257	5 Q9V736	Q9V736 drosophila
37	430	5.3	3111	5 Q9VH10	Q9VH10 drosophila
38	429	5.3	2748	5 Q9V4P4	Q9V4P4 drosophila
39	429	5.3	3261	4 Q9Y556	Q9Y556 homo sapien
40	429	5.3	3664	4 Q9GT58	Q9GT58 homo sapien
41	425	5.3	1110	13 Q91255	Q91255 petromyzon
42	420.5	5.2	4900	5 Q9N541	Q9N541 caenorhabdi
43	419	5.2	3187	11 Q63714	Q63714 rattus norv
44	414	5.1	1422	6 Q95KU4	Q95KU4 canis famil
45	414	5.1	5085	11 Q9JKS6	Q9JKS6 rattus norv

ALIGNMENTS

RESULT 1

Q62766 PRELIMINARY; PRT; 1596 AA.
 ID Q62766
 AC Q62766;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE SSeCKS.
 GN 322.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAINS=FISHER; TISSUE=EMBRYO;
 RX MEDLINE=95257957; PubMed=7739556;
 RA Lin X., Nelson P.J., Frankfort B., Tomblar E., Johnson R.,
 RA Gelman I.H.;
 RT "Isolation and characterization of a novel mitogenic regulatory gene,
 RT 322, which is transcriptionally suppressed in cells transformed by src
 RT and ras.";
 RL Mol. Cell. Biol. 15:2754-2762(1995).
 DR EMBL; U23146; AAA79517.1; -;
 DR InterPro; IPR001573; PfamA anch.
 SQ SEQUENCE 1596 AA; 172498 MW; CF78F2D161D9C7E2 CRC64;

Query Match 100.0%; Score 8073; DB 11; Length 1596;
 Best Local Similarity 100.0%; Pred. No. 1.9e-314;
 Matches 1596; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCAGSSTQRSPEQAGSDTTPSELVLSGHGPAEASGAAGDPADADPATKLPQKNGQLSS 60

Db 1 MCAGSSTQRSPEQAGSDTTPSELVLSGHGPAEASGAAGDPADADPATKLPQKNGQLSS 60

QY 61 VNGVAEQGDVHVQENQEQEEVVDVQRESEDEVREKDRVVEEAAANSTAVEDITKDG 120

Db 61 VNGVAEQGDVHVQENQEQEEVVDVQRESEDEVREKDRVVEEAAANSTAVEDITKDG 120

Db 241 LOAESDQAAEEBAKDEGEKEKEPTKSPSPSPVNSSETTSFKKFFTHGWAGWRKKT 300
Qy 301 FKSKEDDLTAKRKQBAKVDDEBKETEPASEBQEPABDTDOARLSADYKVELPL 360
Db 301 FKSKEDDLTAKRKQBAKVDDEBKETEPASEBQEPABDTDOARLSADYKVELPL 360
Qy 361 EDQVGDLEASSEKCAPLATEVDEKMEAHQEVVAHVHSTVKTETEEOGGGEAGGVV 420
Db 361 EDQVGDLEASSEKCAPLATEVDEKMEAHQEVVAHVHSTVKTETEEOGGGEAGGVV 420
Qy 421 VEGTGESLPPEKLAEPQVQEAEPABELMKSRMCVSGGDHTQLTDLSPSEKTLPKHPE 480
Db 421 VEGTGESLPPEKLAEPQVQEAEPABELMKSRMCVSGGDHTQLTDLSPSEKTLPKHPE 480
Qy 481 GIVSEVEMLSQBRICKVQSGPLKLLPSSGLKKLSGKKQKKGCGGDEBPBYQHHT 540
Db 481 GIVSEVEMLSQBRICKVQSGPLKLLPSSGLKKLSGKKQKKGCGGDEBPBYQHHT 540
Qy 541 SPESADEOKGESSASSPEEPETTCLEKPLEAPOGEAEEGTTSDGKKREGITWASF 600
Db 541 SPESADEOKGESSASSPEEPETTCLEKPLEAPOGEAEEGTTSDGKKREGITWASF 600
Qy 601 KKMVTPKVRVRRESDEKBELEKVKASATLSSTDSTVSEMQDEVKTGEBQKPEEPKRV 660
Db 601 KKMVTPKVRVRRESDEKBELEKVKASATLSSTDSTVSEMQDEVKTGEBQKPEEPKRV 660
Qy 661 DTSVSWEALICVGSKKRAKASSSDDGGPRTLGGDSHRAEASKDKAAGTDAVPASTQ 720
Db 661 DTSVSWEALICVGSKKRAKASSSDDGGPRTLGGDSHRAEASKDKAAGTDAVPASTQ 720
Qy 721 EDOAQGSSSPAGSPSEGEVSTWESFKRLVTPRKKSXSKLEEKAEADSSVQLSTEIE 780
Db 721 EDOAQGSSSPAGSPSEGEVSTWESFKRLVTPRKKSXSKLEEKAEADSSVQLSTEIE 780
Qy 781 PSREESWVSIKKIPGRKKRKGKQKQEAQVDSGPGVNEINDDPNVPAVPLSEYNAVER 840
Db 781 PSREESWVSIKKIPGRKKRKGKQKQEAQVDSGPGVNEINDDPNVPAVPLSEYNAVER 840
Qy 841 EKMEAGNTELPOLLGAVVSEBSKTLVHTVSAVAVIDGTRAVTSVEERSPSWISAVTE 900
Db 841 EKMEAGNTELPOLLGAVVSEBSKTLVHTVSAVAVIDGTRAVTSVEERSPSWISAVTE 900
Qy 901 PLEHTAGEAMPVEEYTEKDIIAEETPVLTOTLPEGKAHDDMTSEVDFTSAAVTATET 960
Db 901 PLEHTAGEAMPVEEYTEKDIIAEETPVLTOTLPEGKAHDDMTSEVDFTSAAVTATET 960
Qy 961 SEALRTEEVTEASGABETTDMSAVSOLTDSPTTTEATPVQEVESGVLDTDEEERQTOA 1020
Db 961 SEALRTEEVTEASGABETTDMSAVSOLTDSPTTTEATPVQEVESGVLDTDEEERQTOA 1020
Qy 1021 ILQAVADKVKESQVPATQTVORTGSKALEKVEEVEDSEVLASEKEKDVMPKGPVQEAG 1080
Db 1021 ILQAVADKVKESQVPATQTVORTGSKALEKVEEVEDSEVLASEKEKDVMPKGPVQEAG 1080
Qy 1081 AEHLAQSETQATPSLEVPETADVHVATQCVIKLQOLMQAVAPSESSETLTDSETN 1140
Db 1081 AEHLAQSETQATPSLEVPETADVHVATQCVIKLQOLMQAVAPSESSETLTDSETN 1140
Qy 1141 GSTPLADSDTADGTQDETIDSDSKATAAVROSQVTEBAATAQKEEPTLNNVPAQE 1200
Db 1141 GSTPLADSDTADGTQDETIDSDSKATAAVROSQVTEBAATAQKEEPTLNNVPAQE 1200
Qy 1201 EHGEPRDVLPTQBELTAAAVPLAKTEVGOGEVDMLDGKVEEKEQEVFVHSGPNSQ 1260
Db 1201 EHGEPRDVLPTQBELTAAAVPLAKTEVGOGEVDMLDGKVEEKEQEVFVHSGPNSQ 1260
Qy 1261 KAADVTDSEVMGACQEKESTEVOSLSLEEGEMETDVEKEKRETKPQVSEEGQETA 1320
Db 1261 KAADVTDSEVMGACQEKESTEVOSLSLEEGEMETDVEKEKRETKPQVSEEGQETA 1320
Qy 1321 APHEGTYGKPVLTLDMPSSERKALGSLGSPSLPDQKAGCIEVQVOSLDTVTQTAE 1380
Db 1321 APHEGTYGKPVLTLDMPSSERKALGSLGSPSLPDQKAGCIEVQVOSLDTVTQTAE 1380

Qy 1381 AVEKVJETVVISETGSPSPCVGAHLLPAKSSATSGHWTLOHAEDTVPLGPSQAASSIPI 1440
Db 1381 AVEKVJETVVISETGSPSPCVGAHLLPAKSSATSGHWTLOHAEDTVPLGPSQAASSIPI 1440
Qy 1441 IYTPAPESTLHPDLOGEISASORSESEEDKPDAGDADGKSTALEKVLKAEPEILELE 1500
Db 1441 IYTPAPESTLHPDLOGEISASORSESEEDKPDAGDADGKSTALEKVLKAEPEILELE 1500
Qy 1501 SKSNKVLNVIIQTAVDQFARTETAPETHAYDSQTVPAQRLDSREPNRCWTMKAAMKH 1560
Db 1501 SKSNKVLNVIIQTAVDQFARTETAPETHAYDSQTVPAQRLDSREPNRCWTMKAAMKH 1560
Qy 1561 PVPQREDILOVTLBWAQPRKCLPRLQKAPVSK 1596
Db 1561 PVPQREDILOVTLBWAQPRKCLPRLQKAPVSK 1596

RESULT 3
Q9WTQ5 PRELIMINARY; PRT; 1684 AA.
AC Q9WTQ5;
DT 01-NOV-1999 (TREMREL. 12, Created)
DT 01-NOV-1999 (TREMREL. 12, Last sequence update)
DT 01-DEC-2001 (TREMREL. 19, Last annotation update)
DE SSECKS (PKC binding protein SSECKS).
GN AKAP12 OR SSECKS.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALE/C; TISSUE=TESTIS;
RA Okita K., Kitamura H., Iwanaga T., Morimatsu M., Saito M.;
RT "Isolation and characterization of mouse SSECKS.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
RN [12]
RP SEQUENCE FROM N.A.
RA Camus A., Mesbah K., Rallu M., Babinet C., Barra J.;
RT "Gene trap insertion into murine SSECKS homologue interrupts one of the transcription units while revealing the existence of a second messenger exclusively expressed in the testis.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AB020886; BAA76894.1; -
DR EMBL; AF326228; AAK16150.1; -
DR MGD; MGI:1932576; Akap12.
DR InterPro; IPR001573; PkinA anch.
SQ SEQUENCE 1684 AA; 180694 MW; E569D55762FCB19E CRC64;

Query Match 80.7%; Score 6512.5; DB 11; Length 1684;
Best Local Similarity 83.9%; Pred. No. 3.8e-252;
Matches 1326; Conservative 72; Mismatches 154; Indels 29; Gaps 11;

Qy 1 MGAGSSTEQRSPEQAGSDTPSELVLSGHGPAEASGAAGDPADADPATKLPQNGQLS 60
Db 1 MGAGSSTEQRSPEQAPAESDTPSELSGHGPAEASGAAGDPADADPATKLPQNGQLS 60
Qy 61 VNGVABOQDVHVHVEENQOEVEEDVDVQRESDREKDRVEEMAANSTAVEDITKDG 120
Db 61 VNGVABOQDVHVHVEESQDQEEVTVEDVQRESBDVKEKDKAEWAASSTVVEDITKDE 120
Qy 121 QBETSEBIIEQIPASNNVEMVQPAESQANDVGFKKVFKFVGFKFTVKDKKNEKSDTVOL 180
Db 121 QBETSEBIIEQIPASNNVEMVQPAESQANDVGFKKVFKFVGFKFTVKDKKNEKSDTVOL 180
Qy 181 LTVKKDEGEASVAGDHQBPSPVETAVGESAKSESELKQSTEQEGLTKQEQSSTEIP 240
Db 181 LTVKKDEGEASVAGDHQBPSPVETAVGESAKSESELKQSTEQEGLTKQEQSSTEIP 240
Qy 241 LOAESDQAAEEBAKDEGEKEKEPTKSPSPSPVNSSETTSFKKFFTHGWAGWRKKT 300
Db 240 LOAESDQAAEEBAKDEGEKEKEPTKSPSPSPVNSSETTSFKKFFTHGWAGWRKKT 300


```
QY 465 LTDLSPKTLPHKEGIVSEVEMLSQERI KVQGSPLKKL PSSSGIKL SGKKQKGR- 523
Db 358 LTDLSPKTLPHKEGIVSEVEMLSQERI KVQGSPLKKL PSSSGIKL SGKKQKGRG 417
QY 524 GGGDEEPGEYQHHTSPESADQKQESSASPEEPETTCLEKGPLEAPQGEABEGT 583
Db 418 GGGDEEPGEYQHHTSPESADQKQESSASPEEPETTCLEKGPSEAPQGEABEGA 477
QY 584 TSQGEKKREGITWASPKQWTPKRVRRPSESDEEELKVKXSATLSDSTVSEWQDE 643
Db 478 TSQGEKKREGITWASPKQWTPKRVRRPSESDEEELKVKXSATLSDSTVSEWQDE 537
QY 644 VKTGEOKEPEEKRRVDTSVSEALICVGSKKRABKASSDDDEGPRTLGGDSHRAE 703
Db 538 VRVGEQRESEERKRRVDTSVSEALICVGSKKRABKASSDDDEGPRTLGGDSHRAE 597
QY 704 ASKDKAGTDAVPASTQEQOAGSSSPAGSPSEGEVSTWESFKRLVTPRKKSXSKL 763
Db 598 ASKDKA--DALPASTQEQOAGSSSPAGSPSEGEVSTWESFKRLVTPRKKSXSKL 655
QY 764 EKAEDSSVQLSTEIPPSREESVSIKFTI PGRKKRAGDKQEQAVVEDSGPVEINEDD 823
Db 656 EERAEDSGAQLASEIPPSREESVSIKFTI PGRKKRAGDKQEQAAVEDSGPVEINEDD 715
QY 824 PNPVAVPLSEYNAVEREKME--ACQNTL PQLLGAVYVSEELSKTLVHTVSVAVIDGT 880
Db 716 PDVPAVPLSEYDAVEREKLEAQRANVELPQLKGAVYVSEELSKTLVHTVSVAVIDGT 775
QY 881 RAVTSVEERSPSMTSASVTPLEHTAGEAMPVEVTEKDIIAETPVLTQTLPEGKDAH 940
Db 776 RAVTSABERSPSMISAMTPELHAGSVATPPGVEVTEKDITAEATPALLAQTLPGGKDAH 835
QY 941 DMVTSVDFTSBAVTAETSEALRTEEVTEASGAETTTDMVSAVSQLTSPDTTEBATP 1000
Db 836 DDIVTSVDFTSBAVTAETTEALRAEELTEASGAETTTDMVSAVSQLTSPDTTEBATP 895
QY 1001 VOEVESGLTDEERQOAILQAVADKVBESQVPATQTVORTGSKALKVBEVEDSE 1060
Db 896 VOEVESGLTDEERQOAVLQAVADKVBESQVPATQTVORTGSKALKVBEVEDSE 955
QY 1061 VLASEKEDVMPKGPVOEAGAEHLAQCSSETGOATPESLEVPEVTDVHVATCOVIKQO 1120
Db 956 VLATEKEKVVPPEGVQEAETHEHLAQCSSETVQATPESLEVPEVTDVHVATCOVIKQO 1015
QY 1121 LMEQAVAPESSETLTDSETNGSTPLADSDTADGTQDQDDETTSDQSKATAAVROSQVTEEE 1180
Db 1016 LMEQAVAPESSETLTDSETNGSTPLADSDTPNGTQDQDDETTSDQSKATAAVROSQVTEEE 1075
QY 1181 AATAQKEEPSTLNNVPAQEGHEEPGRDVLPTQOQBLTAAAVPVLAKEVGOBEVDWL 1240
Db 1076 AAAAQTGEGPST-PSSFPAQBEHREKPGRDVLEPT-QALAAAGAVPILAKAEVGOBEAGQF 1133
QY 1241 DGEKVKE----BOEVFVHSGPNSOKAADVTYDSEVMGVAGCQEKESTEVQSLSLEGEM 1295
Db 1134 DGEKVKGQCQKLEVPVHTGPNSSOKTADTRDSEVMEVARGCQETESNEQSISPEKREM 1193
QY 1296 ETDVEKEKRETKPEQVSEEGEQETAAPHEGTYGKPVLTLDMPSSERKALGSLGGSPL 1355
Db 1194 GTDVEKEETETKTEQASEEHEQETAAPHEGTHPKPVLTDMPHISERKALGSLGGSPL 1253
QY 1356 PDQKAGCIQVQOSLDTTQVTAQAEVKEVETVWISSETGESPECVGAHLLPAEKSSATG 1415
Db 1254 PDQKADKCIQVQOSSTPVTQTTEAVKVEETVATSEMDSELECAGQSLPAEKLSGT 1313
QY 1416 GHTLQHAEDTVPGLPESQAESPIIIVTPAPESTLHPDLOGETSASQRESEEDKPDAG 1475
Db 1314 GYGTQLQGEDTVPQGPSQAESPIIIVTPAPESTLHSDLOQREVSASQKQSDENKPDAG 1373
QY 1476 PDADGKESTAIEKVLKAEPELLEESKNKIVLNVITQAVDQFARTETAPETHAYDSQTO 1535
Db 1374 PDAAGKESAAREKILRAPEPELLEESKNKIVQSVITQAVDQFARTETAPETHASDLQNO 1433
QY 1536 VPACRLDSREPNRCWTQKQKAKMKHPVQPPREDLQV 1571
```

```
Db 1434 VPVWQADSGAQOQMLDK-----DESLQV 1457
RESULT 5
Q9DDN8 PRELIMINARY; PRT; 2471 AA.
AC Q9DDN8;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Gravin-like.
GN GL.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]_
RP SEQUENCE FROM N.A.
RX MEDLINE=21095185; PubMed=11165490;
RA Klingbeil P., Frazzetto G., Bouwmeester T.;
RT "Xgravin-like (Xgl), a novel putative a-kinase anchoring protein
RT (AKAP) expressed during embryonic development in Xenopus.";
RL Mech. Dev. 100:323-326(2001).
DR EMBL; AF308810; AAG45105.1; -
DR InterPro; IPR001573; PkinA anch.
SQ SEQUENCE 2471 AA; 267987 MW; 1DC9FD7D443D1129 CRC64;
Query Match 22.4%; Score 1805; DB 13; Length 2471;
Best Local Similarity 31.5%; Pred. No. 4.5e-64;
Matches 551; Conservative 287; Mismatches 539; Indels 370; Gaps 60;
QY 82 EEVVDVGVQRESEVREKORVEEMAANSTAVEDITKDGQETSEIIEQIPASENNVEEM 141
Db 11 EEVSTNEGPN-----MEVISEATPKTEIKEQGGDEAPEATN--ATESPEST 60
QY 142 VQPAEQANDVGFKKVFKFVGFKFTVKRKNKESDVTQLLVKKDEGEAASVAGADHQ 201
Db 61 EQANETQSNVEGVFKFVGFKFTVKRKNKESDVTQLLVKKDEGEAASVAGADHQ 116
QY 202 EPSVETAVGESASSELKQSTKEQGTLLKQESSTTEIPLQAESDQAAEAKDEGEKQ 261
Db 117 KQN-----DVSGENVLE--TEDTKEVEKQSETDPIIQTSTREEVSGKPKSEETKL 167
QY 262 EKQPTKSPSPSPVNSSETTSFKKFTGAGWRKTSFKKSKEDDLETAEK-----R 315
Db 168 EEEQVTPSPPTNPLVAETSSPLKKFTQGWAGLRKTSFRKSKEDDHQEVKVIISSEK 227
QY 316 KEQAEKVDEEKE-----KTEPASSEEQPAEDTDQARL 349
Db 228 EKSDSQVKDEEDGENQIIVNGETLPKEANEIESKACIDKETQPSAKELAKPEECPTDQ 287
QY 350 SADYEVK-----ELPLEDQVGDLEASSEKCAPLATEVFEKMEAHQEVVAEVH--- 398
Db 288 SVLKEDIKLENADICEVPVTESTI---QTSAAEVHVSVPTEVSEVQKKEVDTTEDNDVKVT 344
QY 399 -VSTVEKTEEEQGGGAEAGGVVVEGTGSLPPEKL--AEQEVVQEAAPAEELMKSR- 453
Db 345 VVSTSTATSEE-----VKEPLSEMAPDTEVICDKNDKDTTKLAQTATLLESTE 393
QY 454 -EMCVSGGDHTQLTDLSPBEKTLPHKEGIVSEVEMLSQERIKVQGSPLKKL PSSSGLK 512
Db 394 VEICL-----EPIMTETELSSQEKAKLQGSPLKLFSGSLR 431
QY 513 KLSGKQKQKRGGGGDEEPGEYQHHTSPESAD--EQKGE--SSASSPEPEETTCLEKG 569
Db 432 KLSGKQKQKRGGGGDEEPGEYQHHTSPESAD--EQKGE--SSASSPEPEETTCLEKG 485
QY 570 PLEAPQDGEAE--EGTTSDEKKEKREGITPWASFKKMTTPKKVRPSPSKDEEELKVKSA 628
Db 486 AEDATQAVETEGDQATSDGERKREGITPWASFKKMTTPKKVRPSPSKDEEELKVKSA 545
```

```

QY 629 TLSTSDTVS-EMQBEVTVGEOK-----PEPKRRVDTSVSWELICVSSSKTARKKS 683
D 546 TWSSTDSAGSVNOBEAKENGEOKLEKSTENKKVDSVSWELICVSSAKKARKKS 605
QY 684 SSSDGGPRTTGGDHRHAEASKDEAGTDAVPASTOBDQOQSSSPBPASPSGEGV 743
D 606 DSDDE-TQKNIDENKKEIEVVKSKHESEDPVSSQKEOVNDIPSPQSTSPTEGDOG 664
QY 744 STWSEFKRLVTPRKKSLSKEKAD-----SSVQLSTEIEPSREBSVWVSIKKFIPGRK 799
D 665 STWQSFKRLVTPRKSRRAEKEETEETAVSNTQPTSDGTGKEGWSFKKLIPGRK 724
QY 800 KRADCKQOATVEDSGP-----VEINEDPNVAVVPLSTYNAVEREKMAQNTBLPOL 855
D 725 KKSDDKQBPATISDTGKSTEGCEAMEDADAVVPLSFDAEQEKRAQOQKEXTLNDV 784
QY 856 G--AYVSEELKTVHTVSVAVIDTRAFTVSEERSPMWISASTYEPLEHTAGAMFVPV 913
D 785 SKDEKTPBVNSEGLIHATTVVGEERAVTSLBDAPSMISANVETIVEQANLTPRVT 844
QY 914 BEVTEKDIIAERTPV---LTQTLPEKDAHDMVTSEVDFTSAAVATETSEALTEEVT 970
D 845 KERISEITVEBAFIFGEVSGVWSG-----NTINEVELTSEALTA-----LEBAI 891
QY 971 PASGAEFTDMVSANSQLTDSPTTBEATPVQEVBSGVLDTEEBRQOALLQANADKVK 1030
D 892 EYSCAEFTTEMISAASQLDSTFTPAEEVTPPEDEGQSQSLDMPKQMDPVDE-ATELT 950
QY 1031 FES-----QVPATQTVORGSKALEKEVE-----EDSEVLASEKEKQVMP-- 1072
D 961 FOSTILSDKIQUNESITSTSAEHDTKVKIEBANICISCKOBEESTIPSEDEIBECTPVL 1010
QY 1073 -----KGPVQAGAEHLAQSGETGQATPESLEVEPTADVHV--ATCQVYKQLQ 1120
D 1011 AERHDPSELKSIVVESNGA-----DSVSV-----VEHTIDAVIHEVKERY 1050
QY 1121 L-MEQ--VAPESSETLDS-----ETN-GSTPLADST--ADGTQDETITDSD 1164
D 1051 LSSDADITVAQVDTKEHPDAVLEVNREBETSGVSEVQDSSTIILDSQADKIVDTAQ 1110
QY 1165 SKATAVAVQSOVTEBEAATAQKEP-----STLPNNVPAQOEHGEE--FG 1207
D 1111 KOSSEITVLSEKKEVBCALMSDHTKTSVHPKCOFEVTSVSTEDDLPVKEQVSDNSISA 1170
QY 1208 RDVLEPTQOELTAAPVVLAKTEVGEQGEVDMLDGEKVEQ----- 1249
D 1171 VKLVVAAEQVTEKSIPTLEQKELSEBSSSTISVSTVHEQVKEALISDQDALPVK 1230
QY 1250 EYVHSGPNSQKADVYDYSEV-----MGVAGCOEKESTEVQSLIS 1289
D 1231 EQVLNDSISAVKLVDVAAEQVTEKSIPTLEQKELSEBSSSISVSTVHEQVKEALIS 1290
QY 1290 LEEGEMDVVEKERETPEQVSEGEQETAPRHEGTVGKVLVLDMPSSRGKALGSL 1349
D 1291 DEQETTERHPHASEVETK-EGAAEQVTEKSLIYEVESQAQVEBCSLFATQK-- 1345
QY 1350 GSGPRLPHOD--KACGIEVQVQSLDTTQVTAQAEVKY-----ETVVISSTGSPCEV 1401
D 1346 -ESVSIIEBELARKSDPVEPLEQS-GITLATYEDTTERKIFADKETATASVAGISAKKA 1403
QY 1402 GAHLPAEK-----SSATGHWI--LOHADTVPLGSPQASIPITVPAPESTILH 1451
D 1404 LNALFENKHSILEDVSEVGKQVSCSAVMPDEQAVEAQAQSTEKQ--VAKSPTPATA- 1460
QY 1452 PDLQGEIASQCHRESEF-----DKPDAGPDAG 1480
D 1461 --LREIASLVEEAQVEEKNAVPSSEFOISIPSSLPIDQKEVTTAEKDTTEIILDE 1518
QY 1481 KESTAI-E-KVLKAPPE-----ILEESKSNKIYLVNQTAVDOFARTETAPETHAYDS 1532
D 1519 TEKQAMESVSVLSGQOEAVNPNFSVVEEVSKAFLAGVEQTJESW-----TTASEE 1569
QY 1533 QTOVPACRLDSREPNRCWTMKDAKMKH---PVPOFREDLQVLTIVLEMAQPRKCLPRL 1588

```

```

D 1570 QTR-----ECTTADVOAVECNATVDEPSEERPSATLSHSAE---TVSV 1613
QY 1589 QLKAPVS 1595
D 1614 SLECPIT 1620

RESULT 6
Q9VC00
AC Q9VC00 PRELIMINARY; PRT; 2768 AA.
AC Q9VC00;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE CG13648 protein.
GN CG13648.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abbill J.F., Ashbyant A., An H.-J., Andrews-Pianko C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokva D., Borchan M.R., Bouck J., Broksrein P., Brottier P.,
RA Butris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong F., Gottrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jaitel M., Kalush F., Karen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusken D.R., Paclob J.M.,
RA Palazzolo M., Platten G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skipski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner C., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AB003750; AAF56376.1; -
DR FlyBase; FBgn0039257; CG13648.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00093; VWC; 1.
DR SMART; SM00214; VWC; 4.
DR PROSITE; PS01208; VWF; 1.
SQ SEQUENCE 2768 AA; 294032 MW; CA929A21774E4684 CRC64;

```

Query Match 8.0%, Score 642; DB 5; Length 2768;

01-MAR-2002 (Tremblrel. 20, Last annotation update)
 DE Erythrocyte membrane-associated giant protein antigen 332.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 NC NCBI_TaxID=5833;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=FCCL/NH;
 RA Shan Z.X., Yu X.B., Li X.R., Ma C.L., Fang J.M., Lu J.H., X J.;
 RT "Molecular cloning and structure analysis of the Plasmodium falciparum
 erythrocyte membrane-associated giant protein Ag332 (Pf332) gene";
 RL Submitted (FEB-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF202180; AAF15293.3;
 DR InterPro; IPR001313; Ptm110/Puf.
 SQ SEQUENCE 5458 AA; 615269 MW; 4F6A37D92BD0172 CRC64;
 Query March 7.9%; Score 635; DB 5; Length 5458;
 Best Local Similarity 23.4%; Pred. No. 5.6e-17;
 Matches 411; Conservative 292; Mismatches 688; Indels 366; Gaps 79;
 QY 4 GSSTQRSPEQPA--GSDTPSELVLGHPAAASGAADPADADPATKLPQKQQLSSV 61
 DB 977 GSVTQLVBEKEGLINENDEKELIT--EMSEIKSVMEIETDLSTELIKQGLATY 1034
 QY 62 NGVAEQ---GDVHVOEN---OEGOE---EVDVEDVQGRSESDVREKORVEEPAANST 111
 DB 1035 EFVEBEKSLTDKLBEESVTKEVGETELSTQEVNDEKVSVEEVEEKSVEEVLBEES 1094
 QY 112 AVEDITKQOQ-----EETSEIIEOIPASE----- 135
 DB 1095 ATEEVEERGSCTEIVEKSGDTEIVEEGSATEIVEEGSATEIVEEGSATEIVEEGSATEIVEEGS 1154
 QY 136 NNEVEVPAESQANDVGFKKVFKFVGFKTYKDKNEKSDTVOLLTVKD-----E 187
 DB 1155 SATTEVEVEGSLSDNVQSKGIENVGELIYSVTKAKDESMN--EKIPLKSFVVDSEFK 1212
 QY 188 GEGASVAGAGD-----HOEPSYETAV-----GESAKSESL--KQ 221
 DB 1213 GGGPDNVNVEDVNSEDIINEHTPLEETKIELPLEYITADITHGTETKYNLYEKI 1272
 QY 222 STKEKGTIKQOSSTEIPLQAS---DQAEBAKADGEGEKO---EKEPKSPESPSS 274
 DB 1273 NEEVEKAKFOEKKITENTIVERESTVEDIVQEPSLAQVEQKESDTNLEETKLANEKII 1332
 QY 275 PNSTSTTSFFKFTTHGAWGRKKTSPFKSKEDDLTETAKRKEQEA---EKVDEEEKK 330
 DB 1333 PEVSTENVNER-----EGDTEBEVLDEDSITEIVEEVSACBEIVEEBS 1382
 QY 331 TEPASEEOPADTQARLSADYEKV--ELPLEDOVGLLASSSEKCAPLAT--EVFPE 385
 DB 1383 ERIVEEESSESEIIVEEESSESEIIVEEESSESEIIVEEESSESEIIVEESTBEVLE 1442
 QY 386 KMEAHQEVAAHVHSTVTE-----EEGGGGG---ABGGVAVVEGSELSPPE 431
 DB 1443 EGSVTEELIVER--EVSTTEBVKIGVSEBEVLBEHGTETPEVQOQSVLEIETESSTE 1501
 QY 432 KLAEP-----QVPOEABP--AEELMKSRMCVSGGDHTQ/LTDLSPREKTLPKHPGIVSE 485
 DB 1502 KPLEDVGSVNEIIVQEGVADQEIYHE--VSTTEKHENVDRSTTEELVER--VQSVSE 1556
 QY 486 ---VEMLSQERIIVQGSPLKTLFSSSGI-----KKSQKQKQKRGSG 526
 DB 1557 EIIVEESASEEIVBEGSVTEEEVEBEKLINEVGATESVTEELIYQEVDAEVLQOEGS 1616
 QY 527 GDEEFGVYHYTE--SPESADQKQSSASSPEEPETTCLEKGLPAPQDEAEGTT 584
 DB 1617 MNEELLEKSIIVEIIVGPEGVTEELVDYGSFAEEVKEEELVTE--EAVQ---YEGSV 1669
 QY 585 SDGKKRKREIT-----PMASFKQMTVPKRVRRPSSD--KEEELKVSAT--LSSTDST 636
 DB 1670 TTEIREEBITENEALIEBSAPAEIIEK---GPNDELIVKEEGDTEELIYNEVSTTDEV 1725
 QY 637 VSE---MODEVKTGVEEQKPEEPKRRVDTSVSWEALICVGS--KRRARKASSSD--D 687

DB 1726 IEEKLVNEQI--VGEESVTEKEVEVERSAT--EDLVEEGASVTEKSVHGGSTTEQLD 1782
 QY 688 ECGPTLQGDSHRAEASKDKAEAGTDAV---PASTOEDQAGSSSPPE-----AGSPS 738
 DB 1783 ESVAEIIVEEVSVDKIIIEEVSVDVEVEBEGSVTEIIVEEBSVPEIILEBELSGSEE 1842
 QY 739 -----EGEGSVTSESPFKRLVTPRRKSKSKLEKADSVQSLTEIIPGRBSW 787
 DB 1843 VLEDEWYTDAMQEGSVTEIEIEEIV--DOEGSTTEIIVDGSNKEIIVEEBSRVEEV 1899
 QY 788 VSIKKFIPIGRKKKADQKQEQATVEDSGPV--EINEDDPNAVVP/PLSYNVRERKMQ 846
 DB 1900 LKKEGFTVTEIIEBGSVTEQ---VEDIKTVSEKSEBSAIEEVEKVEEBSISEKIYKE 1956
 QY 847 GNTLEPQLLAGVAVVSEELSKTLVHTVSAVVDGTRAVSVSEERSPSWISASVTEPLEHTA 906
 DB 1957 ES-----VTEELVRQEESTTEKIVKDVSPTEPFVQ-----TDSVTEKVIQ 1999
 QY 907 GEMPFVEVTEKDIIEETVLTQTLPEGKADHDNMTSEVDFTSEAVTATFSEALRT 966
 DB 2000 GSNTEVADVEKEBSASDE-----HEQBDVSNVQAQVTECKKSVTEIYDEVERT 2048
 QY 967 EEVTEASGAETTDVAVASQLTDSPTTEATPQVEYESGVLDTEEBERQTOALQAVA 1026
 DB 2049 EIIVEENG--VTBEGVDETGSVTE--EIIIEATVTEEVEDGSVTEEVEVEDGSVIOEVE 2104
 QY 1027 D-KVVEEQVAVATQVQRTGSKALEKV-----EEVEDESEVLASEKEDVMPKPVQ 1078
 DB 2105 DGSVTEE-----IVQNGSVTEIIVEEGSVANEVEEVSQSEVDTEVTE--EVEE 2156
 QY 1079 AGA--EHLAQSEGTQATPESLE-----VPEVTADVHVATTCQVIKLOQLMEQAVAPSS 1131
 DB 2157 EGSVVEEIVE--EEGSVVEEIVEEBSVVEEIVEEBSVVEEIVEEBSVVEEIVEEBS 2214
 QY 1132 ETLTSEINGSPPLADSTAGTQODETIDQDSKATAVAQ-----SQVTEEAATQ-- 1185
 DB 2215 VVEEIVEEBSV--VEEIVEEBSVVEEIVEEBSVVEEIVEEBSVVEEIVEEBSVVEE 2273
 QY 1186 --KEEPTLPNNVPAQ-----EEHGEPRGVLEPTQOELTAAVFLAKTE-----VG 1232
 DB 2274 IVEEBSVVEEIVEEBSVVEEIVEEBSVVEEIVEE--GSVVEIIVEEBSVVEEIVE 2331
 QY 1233 QEGVDMLDGEKVEKQEV---FVHSGPNSQKAA--VTYSEVWGVAGCOEKESTEVOS 1287
 DB 2332 BEGSV-----SEVVDTELVNDEIVEQAPFTEVEEVSQSVNDEIIEEDSVAE----- 2378
 QY 1288 LSLERGETTVEKREKRTKPRQVSEEGEOTFAPE---HSGTVKPLVTLDMPSERGX 1344
 DB 2379 -AVEESBITESVQEEETEKGFVLEKVEGTGAVTEEIVQDLITEI--LSESUNGE 2435
 QY 1345 ALGSLGSPSLPDD-----KAGCIEVQVQSLDVTVTQTAEAVKVIETVISETGE 1396
 DB 2436 IINKSDABEILFETFLNEVVGQGSSEIIVEEBSVTEKEVESVTEELV--DEGSV 2494
 QY 1397 SPECVGAHLIPAKESAT-----GGHMTLQHAEDTVP/PGBSQAESPIIYTPAPESTL 1450
 DB 2495 TEELV-----DEGSVTEEVEEGGSIQAEIIVEE-----ESATTEI-----IRDETIV 2536
 QY 1451 HEDDLOGETISASQRESEEDKPADP--DADGKESATRAIKVLAKEEILELESKSKIYVN 1509
 DB 2537 EEVLKEGSA-----TEEIVQDGSQTNFVGQGSVIEEV--EEIISTTEK-----LK 2584
 QY 1510 VIOTAVDOFARTETAPE 1526
 DB 2585 EASALIEEVEBSIRE 2601
 RESULT 8
 ID 09SYM2 PRELIMINARY; PRT, 17352 AA.
 AC 09SYM2;
 DT 01-DEC-2001 (Tremblrel. 19, Created)

DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
 DE I-connectin.
 GN I-CON.
 OS Procambarus clarkii (Red swamp crayfish).
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Astacidea;
 OC Astacoidea; Cambaridae; Procambarus.
 OX NCB TaxID=6728;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21423462; PubMed=11532946;
 RA Fukuzawa A., Shimamura J., Takemori S., Kanzawa N., Yamaguchi M.,
 RA Sun P., Maruyama K., Kimura S.;
 RT "Invertebrate connectin spans as much as 3.5 micrometer in the giant
 RT sarcomeres of crayfish claw muscle.";
 RL EMBO J. 20:4826-4835(2001).
 DR EMBL; AB055861; BAB64297.1; -.
 DR InterPro; IPR000577; FGGY.kin.
 DR InterPro; IPR003961; FN_III.
 DR InterPro; IPR003006; IG_MHC.
 DR InterPro; IPR001452; SH3.
 DR Pfam; PF00047; fn3; 5.
 DR Pfam; PF00018; SH3; 1.
 DR PROSITE; PS00933; FGGY_KINASES_1; UNKNOWN_3.
 DR PROSITE; PS00002; SH3; 1.
 DR SEQUENCE 17352 AA; 1962348 MW; 4BA157BEC042B42D CRC64;
 Query Match 7.3%; Score 588; DB 5; Length 17352;
 Best Local Similarity 22.3%; Pred. No. 1.6e-14;
 Matches 381; Conservative 251; Mismatches 637; Indels 438; Gaps 76;
 QY 41 DPADADPATK-----LPKNGOLSSVN-----GVAEQDVHVOE 74
 DB 14336 DLADERPKTEIAKSLLPQDSIAVQSVKAPVDIKELPKSETAKPVVSTQESLSVE 14395
 QY 75 ENQEQGEVW-DEDVGQRESDVRKDRVEEMAANS-----TAVE 114
 DB 14396 VQVEGQTEWTPKQISETARVSVTKDSQVLEVSSELAPDDIPETRLIEDKAKVTVE 14455
 QY 115 DITKQGEETSEITIOIPASENVEEMVQPAESQ-ANDVGFKVFKFV-GEKFTVKDK 171
 DB 14456 NQQRADQEV--IVLRGPTKEKIIIVDESDEELVKELPEVKVEEIVEEQFKLPKTKK 14513
 QY 172 NEKSDTVQLLTVKDB--GEGAPASVAGDHQPSVETAVGESAKESBLKQSTEQEGT 229
 DB 14514 KVKEETEEVLIKPEKEVEIEVVIKKPDEIEVTEVMIKKPKKKTVEVVEEVT 14573
 QY 230 LKQEQSGSTEIPL--QAESDQAABEAKDEGEKQ-----EKPTKSPSPSPV 276
 DB 14574 IKKPEVIEVTVKKPEKIEEITEEVIKKPEKVVEEITEEVIQKKPKKPEPKIEV 14633
 QY 277 NSETSSFKFFTHGWAGRKTSFKSKXEDDLTAKEKEQAEKVDEE-----EKK 330
 DB 14634 TEEVT-----LKKPEEKIAEVEETEEIIVKKAEEKVEEVEVIQKKPKKK 14679
 QY 331 TEPASEE-----QSPAEDTOQARLSADYKVELPLEDQVGLASSEKCAPLATEVF 383
 DB 14680 PEPKIEEVEEVIKKPEEVEEVIKKPEVVEEVEEVEV--TIKKPEKVVEEVEEVT 14737
 QY 384 DEKMEAH-----QBVAAVHVSVTEKEEGGGEAGGVVVBGTGE---SLPPEKL 433
 DB 14738 QIKKPKKTPPEPKIEEVEEVEE--TLKKPEEK-----VVEEVEEVIKKPKKK 14784
 QY 434 AEP--QEVPOEA---EPAELMK--SREMCVSGDHTQITDLSPEKTIIPKPEGIVSEV 486
 DB 14785 PPKPIEVEEVEEVIKKPEEKIEEVEEVIKKPEKVVEEVS--EVTIKKPEEKIEEVE 14843
 QY 487 EMLSSQERIKVOGSPKLKLFSSSGLKLSGKQKGRGGGDEEPOEQHIHTESPE-SA 545
 DB 14844 -----TEEVKIKRPPKK-----PEPKI 14861

QY 546 DEQKGESSASSPEEP--ETTCLKGLPLAPQDGEABEGTTSDEGKKREGITPWASFMM 603
 DB 14862 EEVTEEVIKKPEEKIEEVT--EEVIRKPEKVVEE-----14897
 QY 604 VTPKRVRRPSDSKEELE--KVKSATLSSTDSTVSEMDQDEKTVTGEQKPEEKRRVD 661
 DB 14898 VSEEVTIKAPKEVVEEVEEVIKMKPKKPEPKIEEVEEVT--TI--KKPEE---KID 14950
 QY 662 TSVSWEALICVGSKKRKARKASSDDDEGPRTLGGDSHRAEASAKDKEAGTDAVPASTOE 721
 DB 14951 EEITEBEVIKRPKPKPEPKIEEVEEVIK-----KPEKVVEEITEEVIKKPE 15001
 QY 722 QQAQGSPPPEAGSPSEGEVSTWESFKRLVTPRKKSUKLEEKAEKSSVQLSTEIP 781
 DB 15002 EKVPEEVEEVIKKPEEKVVEEVEEVIKMKLKKKPEPKIEEVEEVIKKPEEKIDE 15061
 QY 782 SREESWVSIKKIPGRKKRAGDKQEQATVEDSGPVEINEDDPNPVAVPLSEYNAVER- 840
 DB 15062 EITEE-VQIKR---KPKKKPEPKIEEVT-----EVTLRKPKPEAIEITEEVIKKP 15110
 QY 841 -EKMEAQGNTELPQLGAVVSEELSKTLVHTVSVAVIDGTRAVTSVEERSPSWISASVT 899
 DB 15111 EEKIE-----EVSEEVIKKPEKVVEEVEEVIKMKPKKPEPKVEEVESE---EITIK 15162
 QY 900 EPLEHTAGE-----AMPVVEEVT-----KDIABETPVLTLTLPKGDAHDM 943
 DB 15163 KPEEEVAEEVQIRKPKKSVFKTEISEITILQPKPEKEVEEVEEVIKKRPEK-----15216
 QY 944 VTSEVDFTSEATATSETSEALRTEEVEASGAEETDMVSAVSQLTDSPTTTEEATPVE 1003
 DB 15217 IVEEV---SEEVIKKPDEKV-VSEVTE---EVIKMKPKMKPKVEEVEEVEEVT-IKQ 15267
 QY 1004 VBSGVLDTSEERQATIIQAVADVKEE---SQVPATQTVQRTGSKALEKEVEEVEE 1060
 DB 15268 PE-----EVEEIEEVPFLKPKKPKPEKYEEVSEVTIMR---KREDKVEEVEDVH 15318
 QY 1061 VLASEKEDVMPKGPVQEAAGAEHLAQSSTGOATPESLEVPEVTDVHDVATCQVILQ 1120
 DB 15319 IKMKPREK---POEKFDDVSEBITISKPEKQVTE---EITEDVQ-----IKLK- 15362
 QY 1121 LMEQAVAPESSETLTDSETNGTPLADSDATDGTOD--ETID-SQDSKATAVRSQVT 1177
 DB 15363 -----PKKPDITEEVEEVTIIRKEEPEEEDIESETVOIKRKKRKITTEEEVT 15415
 QY 1178 EBEAATAQKEEPTLNNVPAQEEHGEPRDVLPTQOELTAAAVPVLAKTEVQGEV 1237
 DB 15416 E---VTIQPKPK-----EPEVEEVEEFLMKRPPKPKPEEQVEEVEVTI 15459
 QY 1238 DWLDGE-KVKEE--QEVFVHSGPNSQKAADVTYDSEVMGVAGCQEKESTEVQSLSEGE 1294
 DB 15460 QKESPEPTVEEVEEVSFRVKPKKPIVTEIESEEFTRKPKPE-----EEGP 15510
 QY 1295 METDVEKE-KRETGP-----EQVSEGEQETAPE---HEGTGKPVLTLDMPSSERGA 1345
 DB 15511 YEVSVEEVIKKRPPKPIRETQEISEEVEEMKLTPEDEFTHEETEDVKIMKPKPKK- 15566
 QY 1346 LGSLSGSPSLPDQDRAGCIEVQVQSLDITVTOTAAVEKVIETVVISETG--ESPCEVGAH 1404
 DB 15567 -----YSVTE-----ETIDSKWVQESDEEPEEVIYEE 15594
 QY 1405 L-LPAEKSSATGGHWTLOHAEDTVPLGPESQAESIPITVPAPESTLHPDLQGEISASQ 1462
 DB 15595 LEDVVVQKKIIKDG--TEIAQEFVQLRKKPKKSVTV-----15629
 QY 1463 RERSSE-----EDKPDAGP-DADGKESTAI-----EVLKAPPEILESKNKIVL 1508
 DB 15630 RDEGETVVIKSIITDGTAGPKHEDIEESFQLRPPKPKKYKVEEVEEVEGVKIRREPT 15689
 QY 1509 NVIOAVDOQFARTET-APETHAYDSQTVQVACRLDSRPNRCWTWKDKAKMHPVPQPRE 1567
 DB 15690 FEVRDEEVESFRIKTWEE-----EPEEEVEE-ROVEGKFTFDLPK 15731
 QY 1568 DIQVLTVLAWAQRKCLPRLQLKAPV 1594

Db 15732 KIQKRTTEIEYTKVK-LPPLPKQPPV 15757

RESULT 9

Q22248 PRELIMINARY; PRT; 1829 AA.

AC Q22248; 01-NOV-1996 (Tremblrel. 01, Created)

DT 01-JAN-1999 (Tremblrel. 09, Last sequence update)

DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)

DE T06D8.1 protein.

GN T06D8.1.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

OC Rhabditidae; Pelodierinae; Caenorhabditis.

OX NCBI_TaxID=6239;

RN NCBI_SEQUENCE FROM N.A.

RA Palmer S.;

RL Submitted (Apr-1995) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RE MEDLINE=99069613; PubMed=9851916;

RT none;

RT "Genome sequence of the nematode C.elegans: A platform for

RT investigating biology."

RT Science 282:2012-2018 (1998).

DR EMBL: 249130; CA88964.1; -

DR InterPro; IPR003014; PAN.

DR Pfam; PF00024; PAN; 3.

SO SEQUENCE 1829 AA; 187218 MW; 18A4564A71B3C3E CRC64;

Query Match 7.3%; Score 585.5; DB 5; Length 1829;

Best Local Similarity 23.2%; Pred. No. 1.5e-15;

Matches 375; Conservative 174; Mismatches 662; Indels 405; Gaps 65;

QY 2 GAGSSTE-----QRSPQAGSDTPSELVLSGHGPAASG-----AAGDPADADPAT-- 49

Db 106 GSGEETTVAAVSSGSEEPASSSTSVPTLSKDDQVTEASGEETTTAAATEASSETTS 165

QY 50 KLPQKNGQILSVNGVAE-----QGDVHVOENQGGEE-----VYDDVQ 91

Db 166 AYTEGSGEETTTSAVTEASSEATTPAGTEASGEETTTSAVTEGSEETTVAAVSSGE 225

QY 92 RSE-----DYREKORVEEMANSTAVEDITKDGQETSEIIEQIPASENNVEEMV-- 142

Db 226 EPASSSTSIPTLSKDDQVTEASGEETTTAAATEASEETTTSAVTEGSGEDTTVAVVEL 285

QY 143 ---QPAESQANDVGFKVFKFTVKKDKNEKSDTVQLTVKKDEGEGASVAG- 198

Db 286 SGEQPASS-----TSIPTLSKDDQV-----TEASGEETTTAAAT 321

QY 199 DHOEPEVETAVGESASKESEIKOSTEKQSTGLKQBOSSTEIPIQASDDAEEADEGE 258

Db 322 EASEETTTAVTEGSEETTVAAVSSGE--EPASSSTSIPTLSKDDQVTEAS--GE 376

QY 259 EKOQKEPTK-SPESSPVP---NSFTTSFKKFTTGMAGMRKTKTSFKSKEDDLETAEK 314

Db 377 ETTTAAATEASEETTTSAVTEGSGEDTTVAVSSGEQPASSSTSIPTLSKDDQVTEA 436

QY 315 RKEQAEKVDEEKEKTEPASEQEPADT-----DQALADYKEVPLEDOV 364

Db 437 SGEETTTAAATEASEETTTSAVTEGSGEDTTVAVSSGEQPASSSTSIPTLSKDDQV 496

QY 365 GDLEASSSEKCAPLATEVPEDEKAEAOEVAAVHVSTVEKTEEEQGGGAGAGGVVEGI 424

Db 497 --TEASGEETTTAAATEASEE-----TTTSAVTE--GSGEETTVAAVSS 538

QY 425 GSLPEPEKLABPEVPEQAEPAEELMKSRMCVSGDHT---QUTLSPPEKTLPGHPEG 481

Db 539 GE---EPASSSTSIPT-----TELSKDDQVTEASGEETTTAAATDASSSETTTSAVTEG 588

QY 482 IYSEVEMLSQGERIKVQSGPLKULFSSSGLKKGKQKGRGGGDEPEGEYO-HITTE 540

Db 589 SGEETTVAAVE-----SSDEEPASSSTSIPT 616

QY 541 SPESADEQKGESS-----ASSPEPEETTCLEKPLEAPODGEAEGETT-----SDGEK 589

Db 617 L--SKDDQVTEASGEETTTAAATEASEETT-----TSVTEGSEETTVAAVSSGEE 668

QY 590 KREGITTPASFVKQVTPPKRVRPPESDKEELEKVKATLSIDSTVSEMODEVKTGVE 649

Db 669 ---PASSSTSIPT-----ELSKDDQVTEASGEETTTAAATDASSSETTTSAVTEGS 716

QY 650 -----EQPEPEPKRVDTSVMEALLCVGSSKKRKAASSDDEGPRTLGSGSH 699

Db 717 GEETTVAAVSSSDEEPASS-STSIPTL-----SKDDQVTEASGEETT 759

QY 700 RAEEASKDEKAGTDVAPASTOEDQAOQ---SSPEPAGSPSEGBGVSTWSEFKRLVTPR 756

Db 760 TAAATEASEETTTSAVTEGSEETTVAAVSSGSEEPASSSTTS-----IPT 805

QY 757 KKSKELEKKAEDSSVQLSTEIEPSREESVNSIKKIPGRKKRADGKQOATVEDSGP 816

Db 806 ELKSD--DQVTEASGEETTTAAATEASEETTS-----AVTEGSG- 843

QY 817 VEINEDDPNVPAVPLSEYNAVEREKMAQGNTELPOLGAVYSEELKTLVHTVSAV 876

Db 844 ---EDTTVAVSS-----GEOPASSSTSIPT-----TELSK----- 872

QY 877 IDGTRAVTSVEERSPSWISASVTEPLEHTAGAMPVEEYKDIIEETPVLTOGLEG 936

Db 873 -DDQVTEASGEETT-----TAAATEASEETTS-----VTEGS--GEETTVAAVSSG 919

QY 937 KD-----AHDVTSSEVDFTSEAVTATE-TSEALRTEVETASGAETTTDWS 983

Db 920 EPASSSTSIPTLSKDDQVTEASGEETTTAAATEASEETTTSAVTEGSEETTTSAV 979

QY 984 AVSQTLDSEDTTEKTPQVQVBSGLDTEEBERQALQAVADKVK---EESQVPATQ 1039

Db 980 EGS-----GEETTSVAP-----EGENSTTEAPAVTGSIEIIPSSSESSSTTH 1024

QY 1040 -----TVORTGKALEKVEEVEDSEVLASEKEDVMPKQPOVEAGAEH-----LAQ 1086

Db 1025 DPSIVITPKPEVSSSTIERWMS-KTSSEBAEKKIIGHQTKDDADKEDEDMNPAPVT 1083

QY 1087 GSETQOATPESLEVEPVADVDHVAATCOVIKIQULMEQAVAPESSETLTDSETNGSPLA 1146

Db 1084 ANPAGSTTESAEVNTSGEED-----ENIKWAKELGKQPAD-----LA 1123

QY 1147 DSDTADGTVQOEDTISQSKATPAARQOQVTEEAATQKEEPSTLPNNVPAQOEHGEPE 1206

Db 1124 KLAARDGVNLTEETAADAKSGETAHVEDQVSTSESSISGEETTTTV--NKETTEHHENS 1181

QY 1207 GRDVLPTQOELTAAAVPLAKTEVGOGEVDMLGEKVEKOEQVFAVSGPNSQKADVT 1266

Db 1182 G-----EEDDAPAFYTGAFPTDSTTEASVSTTA--ITDETTVAADDESTSAGEVQ 1231

QY 1267 YDSEWAGVA--GCOKESTEVOSLSLEGEMETDVEKEKRETKRPQVSEGEQETPAPEH 1324

Db 1232 SSSAIIISATVASEBOTSEATSVIESSGE--EVTTDENLVSTVAOLEGSGGATVAAESK 1290

QY 1325 EGTGKPVLTLMPSSEKKGALGSLGSPS--LPQODRAGCIEVQVQSLD-----TTVQY 1378

Db 1291 D-----EDSVTTEATSQSTTVSSSSDGSGETVAPADSESTSSQSSTDESGVTAES 1346

QY 1379 AEAVEKVIE--TVVISGEGSEPCVGAHLLPAEKSSATGCHTLOHAEDTV---PIGPES 1433

Db 1347 KDEESSTTEAPAFVTSKTSGSE-----DEEDSPDTEFLTGIDETMKNKSLVPT 1397

QY 1434 QAESIPITVTPAESSTLHPDLQGEISASQRESEEBEDKPADPADGKESTAIEKY 1489

Db 1398 HREDLPNNVGFPVPSSEPKRKNPDE-----EEEBEEDOTKSDVDVDDVNSKKI 1444

RESULT 10

```

076891
ID 076891 PRELIMINARY; PRT; 5327 AA.
AC 076891;
DT 01-NOV-1998 (TEMBLrel. 08, Created)
DT 01-NOV-1998 (TEMBLrel. 08, Last sequence update)
DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)
DE EG:49E4.1 protein.
GN FUTSCH OR EG:49E4.1 OR CG3064.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID:7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Papagiannakis G., Spanos L., Siden-Kiamos I., Louis C.;
RT "Sequencing the distal x chromosome of Drosophila melanogaster.";
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Benos P.;
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL031128; CAA20006.1; -.
DR FlyBase; FBgn0015390; futsch.
SQ SEQUENCE 5327 AA; 575942 MW; FEFEE23A118FF38A CRC64;

Query Match 7.1%; Score 574; DB 5; Length 5327;
Best Local Similarity 21.6%; Pred. No. 1.5e-14;
Matches 415; Conservative 285; Mismatches 734; Indels 484; Gaps 82;

QY 6 STQRSPQAGSDTSELVLSHGPAEASGAAGPAD-----ADPATKLPQKNGQL 58
DB 2744 SVVSESSKDDAEKESRPESVIAEGEPVPRRESKGPLSDSKDTSRPGSVVSVTADEKSEQ 2803
QY 59 SSVNGVAE--QGVHVQENQEQEEVVDVQGESEDVREKDRVENA-----ANSTA 112
DB 2804 SRRESVAESVADTKDGKQSEASRPSSVDLL---KDDKEQESRRQSIGSHKAWSTM 2860
QY 113 VEDITKGQSETSEIIEQIPASNNVEMVQ-----PAESQANDV--- 152
DB 2861 GDESPMDKADKSE----PSRPESVAESIKHENTKDEESPLGSRDRSVAESIKSDITKG 2915
QY 153 -----GFKVFKFVGKFTVKDKNEKSDTVQLLTVKKDEGEAEASVGADHOEP---- 203
DB 2916 EKSPPLPSKEVSRPESVVGSIKDKAESRRRESVAESVKPESKDATSAPPSKEHSRPESVL 2975
QY 204 -----SVETAVGESAKSESELKQSTE-----KQEGTLKQEOSSTEIPL 241
DB 2976 GSKLDGDKTTTTRRVSVADSIKDEKLLVQSEASRPSEAEISLKDAAAPSOETSRPESVT 3035
QY 242 QABSD-----QAAEEAEKDEGEKEKPTKSPESPSPVNSSETTSSPKKFF 288
DB 3036 ESVDKDGSPVASKEASRPASVAENAKDSADEKQRPESLPQSKAGSIKDEKSP----- 3089
QY 289 THGAWGRKKTSPFKSKEDD-----LETAERKKEQ----- 318
DB 3090 -----LASKDEAEKSEESRRRESVAEQPLVPSKEVSRPASVAESVKDEAEKSEESPLM 3143
QY 319 -----FAEKVDEEK-----EKTE-PASEEQPAEDTOARLSADYEKV 356
DB 3144 SKEASRPASVAGSVKDEAEKSEESRRRESVAESKPLPSKEASRPASVAESVKDEADKSK 3203
QY 357 ELPLEDQVGDLEASBEKCAP--LATEVFDE---KWEAHQEVVAHV----- 398
DB 3204 ESRRSEGAESKPLASKEASRPASVAESIKDEAEKSEESRRRESVAESKPLPSKEASRPRTS 3263
QY 399 -----VSTVEKTEEGGGGEAGGVVVEGTGSLPPEKLAEPQVPOEAPAEELMKSR 453
DB 3264 VAKSVKDEAEKSEESRRSDSAESKPL--ASKASRPASVAE--SVQDEAEKSK--ESR 3317
QY 454 EMCVSGGDHTQTLDSLPEEKTLPKH-PEGIVSEVEMLSQERIK--VQGSPLKPLFSS-- 508

```

```

QY 1466 LPAKSSATNGHWTQHAEADTVPLGESQAESIPITVPAPESTLHPDLQGHISASORE- 1464
DB 4372 FESVSTSVTKS--TVLSSOSTVQLREESTESBL-----SSSLKVE-----DSRRRS 4416
QY 1465 -RSEEDKPDADPDGKSTAIKVLKAEPILELESKNKIIVNVIOTAVDQAFARTE- 1522
DB 4417 LSSLLAEKGIINTNSLKEDTS-ASASQBELLVQSESSSSISYSEIOTISIAQSKNEI 4475
QY 1523 -TAPETHAVDSQTVQVACRLDSREPNCWTCKMDAKMKHPVQPREDDLOVTLVLEAMA 1579
DB 4476 KDARET-----KVTSQFTTTTSSATKCDLSLKETVLEAFLEATERKIVSAKEAFS 4521

RESULT 11
Q9FND5 PRELIMINARY; PRT; 2910 AA.
ID Q9FND5
AC Q9FND5
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Similarity to heat shock protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eustoids II; Brassicales; Brassicaceae; Arabidopsis.
OC NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=98069011; PubMed=9405937;
RA Kocani H., Nakamura Y., Sato S., Kaneko T., Asamizu E., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. II.
RT Sequence features of the regions of 1,044,062 bp covered by thirteen
RT physically assigned P1 clones.";
RL DNA Res. 4:291-300(1997).
DR EMBL; AB006702; BAB11602.1; -.
DR EMBL; AB017062; BAB11602.1; JOINED.
DR Heat shock.
SQ SEQUENCE 2910 AA; 325351 MW; A847EC3FE1427DF7 CRC64;

Query Match 7.0%; Score 564.5; DB 10; Length 2910;
Best Local Similarity 21.6%; Pred. No. 1.8e-14;
Matches 404; Conservative 275; Mismatches 645; Indels 549; Gaps 84;

QY 65 AEOG--DVHVOENOG--QEEVEVDEDVGORES-----EDVREK--DRVEEMANSTRA 112
DB 319 AEOGSEVTVDEKEEDITQNIIEVQESPVSVMESTPTIOGEDIESKASLDHEEM----- 372
QY 113 VEDITKQGEETSEIIEQIPA--SENNVEE-----MVQPAESQANDVGFKKVFKVGF 163
DB 373 -DKITQDEQEHVAVRDVPPQSEFTLVTEAKTAEFTFSVQEAAILKTINISESAHSAIG 431
QY 164 K-----FTVKQK--NEKSDTVQLTVKKDEGEAGASVGA--GDHQE----- 202
DB 432 EEDQGTENTEPSPKDLKDKQDESEITVKTI--ISSDEVRSDDVQAEVGEHTEPCSSSI 490
QY 203 -----PSVETAV--GESASKESLQSTOKEOGTLKQOESSTEIPQ 242
DB 491 KDDRHRDSEIEVKAKETGEPEITAVDGSVNH-----ITTEVVLLAEKEBEKEIKID 546
QY 243 AEDDAAEAEANDG-----EKO--EKEPTKSPESPSPVNSETTSPFKKFTTHGMAGM 295
DB 547 EEPFLAIEKAETENVKIVIEPEIYNNSETSVHESESLEKNAEPEAVAYN--SDGTQI 604
QY 256 RKTSTKSKSKEDL-----ETAEKKEQA-----EKVDE----- 325
DB 605 SKEVTVDRKKEEDIAKTEIEIOPESPESKASLEPKKEVHISNTEHEHVLERDVOQE 664
QY 326 -----EKEKTEPA--SEEOEPAE--DTDOARLSADYEKVLPLEPDQVGLDEASSEK 374
DB 665 TTESAVETKEOTQPSLDLKEDEKETEAEETFKVIVSSDEVRSASAVQEBQF-----EH 717

```

```

QY 375 CAPLATEVEFDE---KMEA-----HOEVVAEVH-----VSTVEKTEEBQGGGAEAGVVV 421
DB 718 TEPCSSIEKDESHGKESEVEVKSQETVODENTEDKHVLKVPSTESEKYQGNBE--TVLV 776
QY 422 EGTG----- 425
DB 777 SNTGSYEKSEKSPDLVAVNDKELANDKINIVDQVGTQIMEEPIGLDSNGAEAEQIDQN 836
QY 426 -----ESLPEPKLAEPOEVPOA-----EPAELMKSR 453
DB 837 ITNETEELVAKPVSLIDPVKSVQMQKPLESPSEVSESTSKTVDEKIEKPEEEVTLVYQ 896
QY 454 EMCVSGDHTQULTDLSPEEKLPHKPEGIVEVEMTLSSQERIKVQGSGLK--LFSSSL 511
DB 897 EGQVDS-----YGLETEETV--SVPESI--ELEEQPOEBSVIDPTLPLOKPTLSPSEV 948
QY 512 KKLSGKQKGRGGGDEPEEGYHIHTESPESADEQK-----ESSASSPEEPEET 564
DB 949 LEBSKV-----DEKIEKTDSELGELIAQESVTDLPLOESSQPNQEKET 999
QY 565 CLKK--GPLEAPQDGAEEGTTSDGKKREGITPVASFKNVTPKRVRRSPESDKBEL 622
DB 1000 KLEKHEPTNEVKSDEVIEVLASPSKLEGET-----VVEAKENIKENEEQQA 1051
QY 623 EKVKA-----TLSTDSV--SEMQDEVKYGEEQKPEEKRRV-----DTSVSW 666
DB 1052 EKIQKSLTQVQTESPSSILFSSSEQDHV--TVAEIIVDEKAKEEVPMLQINEDATYIH 1110
QY 667 EALICVSGSKKRAKASSDDEGGPRTLGDSHRAERASKKKEAGTDAVPASTORODAO 726
DB 1111 ETRV-----EAKRIGSLTB-----ICINQOPEQVKE-----ACKSEQKE 1151
QY 727 GSSSE-----PAGSPBEGEVSTWSEPKRLVTPRKKSCLKLEAKDSVEQLST 777
DB 1152 ISTNENINVTYALHSVEAAEEETATNGESLIDVET--TKSVILVRKEEBAEMKT 1207
QY 778 EIEPS-----RESWVSIKKFLPGRKKRADG--QEQATVEDSGPVE--INEDPNVAV 829
DB 1208 DAEPLDAIEKEELETVTV-----QDAIVNNEETTAHESSELGDNHQENAEAPV 1260
QY 830 VPLSEYNAVEREKMEAOQNTLPLQLAGAVVSELSK-----TLVHTSVAVIDGTRAVT 884
DB 1261 EATQMLDABQISREVTYDTE-----READITKEIKEQEGTVLETFTIQED--IESET 1314
QY 885 SYEERSPSMISASVTEPLEHTAGAMPVE-----EYTEKDI 921
DB 1315 SLELKEEVDQSSKQTEHEHVLERDIPQCEITLKAAVADTVTEBAAILKTLETNISPEBA 1374
QY 922 IAEETPVULTPLPEKDAHDDMTSEVDFTSEAVAT-----ETSEALRTEEVTEASG 974
DB 1375 MASETSLDKVDKEQKEAB--TVKTVIFSNVEGTSDAQAEFGHTEPCSEIIDEQSG 1431
QY 975 AETTDM-----VSASVOLTDSPTTEAT 999
DB 1432 SEESVEVSKETVQGESSEBKDVNMLDVQSGSEKYOENEPDILSVSK--TEKGDFEELP 1490
QY 1000 PVQE-----VESGV--LDTEEBEROTQALLOAVADKVKVESQVATQTV 1041
DB 1491 SVVEGAGLDETTNQTLLDVESVVKQSLDTPBEETSTXTIDKIDKPEB--VTLHQEG 1548
QY 1042 QRTGSKALEKVEV-----BEDSEVLASEKEKDVMPGVPQVQACAEHLAQOS 1088
DB 1549 REEGSYGDTQDEAVSVLESRELGEOPQOEBELCLANBENETKQO---EEQVDKHEPTKE 1605
QY 1089 ETG--QATPESLEVEPVATDVHATCOVITKLQQLMEQAVAPESSETLLDSTNSSTPLA 1146
DB 1606 EVSNQOSP-----VEISNEVLYQVSSASI-----SEGE--YETVYEAKKIGBOYA 1651
QY 1147 DSDTADGTQODETTIDQDSKATPAAVRQSQVTEEAATQAKPESTLPNNVPAQEBHG-- 1204
DB 1652 DK-IQSPFETGIVAHNSHSLPSSSEKKEHETVSEKTTDEK-----VKQAEPIGDWR 1701

```

QY 1205 EPCRDVLEPT-----QOELTAAAVPVL-----KTEVQGEVDWLDGEKVEE 1248
Db 1702 ERGLDIAETHTLSLPSVDQEDVDEIHIPSVALLPLDQEKVTSTKGETKSSEAEDDKPD 1761
QY 1249 QEVFVHSGPN-SQKAADVTDSEVMGVAGCQEKESTEVQSLSLEEGSMETDVEKRETK 1307
Db 1762 EHVDSSTSPMLSEKNDNETQS-----KTSADVCMQOESGLTLEVPKEESKEDK 1811
QY 1308 PEQVSEGEGETAAPHHEGYTKPVLTLDMPSRGRKALGSL-----GGSP-----SLPD 1357
Db 1812 SOBISETIEBIATSDQ-----TLPIETSHDNTLSSELVSEODQSPKKVEBIHE 1862
QY 1358 QDRAGCIEVQVSLDITVTOTAAVEKVIETVISETGE-----SPECVGNHLLPAEKSS 1412
Db 1863 EEPKEAHVDATSERNLPLVTSOA-DNTLSQLVSETKEBHKLOAGLITTEIIPRESSD 1921
QY 1413 ATGCHWTLOHAEDTVPLGPSQAESIPITVPAPESTLHPDLOCEISAS-QRERSEBEDK 1471
Db 1922 EALVMSLASREDDKVALQEDCADDV-----RETN---DIQERSISVETESVGETK 1971
QY 1472 P-----DAGPDADQKESHTAEKVLKAEPEILELSKSNKIVNLVIQTAV-----DQFARTE 1522
Db 1972 PKEHEIDARHVEPTAPILLENDSUSETLIAEAKGNEEINERTVALDHDEEFVNH 2031
QY 1523 TAPETHAYDSQTO 1535
Db 2032 APKLEETKDEKSQ 2044

RESULT 12

Q9NFS3 PRELIMINARY; PRT; 16215 AA.
AC Q9NFS3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE D-titin.
GN SLS OR D-TITIN OR CG1915.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]

SEQUENCE FROM N.A.

RP Zhang Y.Q., Broadie K.S.;
RT "Characterization of Drosophila D-titin gene";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
CC !- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
DR EMBL; AJ271740; CAB93524.1; -.
DR HSSP; P56276; 1TLK.
DR FlyBase; FBgn0003432; sls.
DR InterPro; IPR002106; AATRNA_ligasell.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR003598; Ig_C2.
DR InterPro; IPR003600; Ig_Like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001452; SH3.
DR Pfam; PF00041; fn3; 5.
DR Pfam; PF00047; ig; 50.
DR Pfam; PF00018; SH3; 1.
DR SMART; SM00408; IGC2; 15.
DR SMART; SM00410; IG_Like; 34.
DR SMART; SM00326; SH3; 1.
DR PROSITE; PS00179; AA_TRNA_LICASE_II_1; UNKNOWN_1.
DR PROSITE; PS50002; SH3; 1.
KW Immunoglobulin domain; SH3 domain.
SQ SEQUENCE 16215 AA; 1841509 MW; 242C8765E00F7603 CRC64;

Query Match

6.9%; Score 558; DB 5; Length 16215;

Best Local Similarity 20.5%; Pred. No. 2.3e-13;

Matches 390; Conservative 303; Mismatches 720; Indels 490; Gaps 79;

QY 7 TEORSPEQAG-----SDTP-SELVLSHGCPAAEASGAAGDPADADPATKLPQK 54
Db 12422 TTKRMLRRPAGGEIIBIEVRRDQPEABITIVVEYEP-----EPVNDQEKPEPK 12472
QY 55 NGQLSSVNGVAEOGDVH--VQ-----EENQOGEQEBVDE-----DVGORE 93
Db 12473 KTR-----KVKDDIHDYIOKLIETPKTELEKYEKIEFPIVKDKPLDSIDVLDES 12526
QY 94 SEDVREKDRVEMAANSTAVEDITKDGQETSIIIOIPASENNVEEMVQPAESQANDVG 153
Db 12527 PKEVQKDKKSRITCKQCVEEAPQEPPIPVQILEVKPV-EVDVKEVITDCDKPVQEK 12585
QY 154 FKVPFVVG-----PKFTVKDKNKSQVTLTKKDEGECAEASVAGDHOEVSVE 207
Db 12586 TRKVLKGIPEEQTTFKITM--IESEDNSV--TVIVDEEPIASQSTEEHPEOSKEK 12640
QY 208 AVGESAS-----KESL-----KOSTEKQEGTLKQEQSSTETPLQAESDOAA 249
Db 12641 LAPKPKTKVRKVKDDLSYVVKLIEEIPKVDLEKY---KVMEKPKVKLTV-SDSIP 12696
QY 250 EEEAKDEGE-----EKQEKPTKSPSPSPVNSSETTSFKKFFTHGWMGRKKTSPFK 303
Db 12697 EEPKPKSQPISVLPDPTTKPKTKTKPTKPTEDTDOQVPDEPTTVDITDIPELTPTQT 12756
QY 304 SKEDDLETAERKQEAKEVDEEE-----KEKTEPASEEQEPAEDTDQARLSADY 353
Db 12757 AQPEDTATAQITPSAQEEKSTODDTKDTIKQTVKHKTKP--DTQKSVETSELPEVHKDY 12814
QY 354 -----EKVELPLEDOVGDLEASSEKCAPLATEVEDE----- 385
Db 12815 QISIHEELVEEQPEKILEVRVID-EVAEVEESQPIVEVEDEEPEQATEETVEDTKP 12873
QY 386 -----KMEAHQEVVVAHVSTVEKTEBEGGGGGAEGGVVVEGTGESLPPKLAEE 435
Db 12874 KSKKKVKKVKKDDHDLIKKMLEQIEKTELEKYEKIEFDPVKLKPFEAALPIKIER 12933
QY 436 PQEVP-----QEAEPABEL-----MKREMCVSGGDHTOLT 466
Db 12934 KEQKPTKVTILDATDVPKTKLPSKRKEKPABELTVQLPKFLKARMVLVE----- 12985
QY 467 DLSPEEKLTPKPE-CIVSEVEMLS-----SQERIKVQSPLKLFSSSGLKSLGKKQK 520
Db 12986 --YPPAPLIPKTTDIGAINDGELSNIAGEABEILFKPHKTKKI-----KKIK 13032
QY 521 GKRGGGGDEBEPGEYQIHITESADEQKGESSAPPEPEETTCLKGLPEAPQGEAE 580
Db 13033 -----DDLEKVELEKYEKVISSEEPKPYKK-PEKAPKEEKQ 13072
QY 581 EGT---TSDGEKK-REGITPWASFKQKVTKKRVRRPSBSDKB-----EELEKVKSAT 639
Db 13073 EDVKLKLGLGKKKKPEEAP-----ENVTLKNIPQKPEVEEVELKQKPEVEIVEEQ 13127
QY 630 LSSTDSTV-----SE-----MQDEVKTGVEEQPEEPKRRVDTSVSWEALICVSSSK 676
Db 13128 KKPQGEFVVVEPEPSEFDRPEYVPDELEQIEHPKPEIKVKKPSKT-----K 13174
QY 677 KRARKASSDDGGPRTLGGDSHRAEASKDKAEAGTDAVPASTQEQDQAGSSSPBAPS 736
Db 13175 YKPKDKSKSEPETIVSEIVAGVPKEEALPQODVKRKPERDAPEETDSEIKLRPVQAS 13234
QY 737 PSEGEQVSWESFKRLVTPRKSKSKLEKAEADSSVE-----QLSTIEPSREESWVS 789
Db 13235 KDN-----PDEQALVTP--KAEKPIQIEEDKADDEKPKKSKPKKVPQKPEQ--- 13282
QY 790 IKKFIPIRRKKRADGKQEQATVEDSGPVINE-DDPNV-----PAVPLSEYNAVEREKME 844
Db 13283 IAKBED--EEFEVSKEEALVDK--PIEIKPKDKVKKKPEKPEAPSEVVVIEEKP 13338
QY 845 AQNTLEP---QLLGAVYVSELSKTLVHTVSVAVIDGTRAVTS--VEER----- 889
Db 13339 EEPSEIPVEYKITTVTLPEPEDAPKE--HGVKVIDFERQETTEEVEIEKVKKKPKP 13396
QY 890 -SPSWISASVTEPLEH-----TAGAMPVPVEVTEKDIIAETPVLTTQTLPEGDAHDD 942

```

Db 13397 QPPEFEVTLKPKPEQIOIPDVSAISLPIEBPECKPPOYEVELKITOTTER--EPNUV 13454
QY 943 WVTSEVDFTSVAUTATETSEALRTBEVTEASGAETTDVMSAVSQ--LTDSPTTE-EAT 999
Db 13455 QLAIVEKVTKVKKVKKEDKIVVEAEKEKOPVEETIVEVEKQBEKKSEKSPKSYFKIS 13514
QY 1000 PVOQVESGVLDTEEE--EKQTAIILQAVDK-----VKE-----EQVRA 1037
Db 13515 ETQSIIEKKIEVAEAEAEETPKVKEKVAEKDYSFTLKEETDEEKVITVDDQPEEAP- 13573
QY 1038 TOTVORTGSKALEKVE-----VEEDSEVLASEKEDVMPKQVVOAGAEHLAQSS 1088
Db 13574 VEVVKKKKKEEBAVAEAEVMTPEKIVETSVETAKQKKTTPKDBEBAQAIIVVSS 13633
QY 1089 ETGQATPESLEVEPVYADVAVATCOVILQOLMEQAVAPESSETLT---DSEINGSTP 1144
Db 13634 EAPVAEEVSEAPE-----SKIVE-----EVIAEKKPEFTINVSESEPKPEBP 13678
QY 1145 LADSDTADGTQODET-----ISQDSKATRAVQ-----SOYTEEATATQK 1186
Db 13679 SVEQFTVKKRKKSPVTFADPATEIVKESKPAEVVTEDAHIKTKKKKVTQVEAEELKI 13738
QY 1187 BEPSTLPNNVPAOEHGEHPGRDVLPTQOELTAAVPIAAKT-EVQGEVDMLDGKRY 1245
Db 13739 KITTEVPQETIPIEEVSE--EVITETKK---TAPVBEKTYKIGIKETPEKPAEKI 13791
QY 1246 KEEQEVFVHSGNSQKAVDVTYDSEVGVAGCOEKESTEVOSLSLEGEMETDVEKEKE 1305
Db 13792 VEEBEVTVTEPIEAPKPEVPEHKVRVI-----EETPRELVEEVLIEEVKVIIRK 13842
QY 1306 TYPEOVSEGEOTAPAEHEGTYGKAVLTLDMPSSRGALGSLGSSPSLPDODKAGCIE 1365
Db 13843 KKPKEIKERPEABVTV---STPKRV-----BEVATSSIAVIPEOPTEEEPADAK 13889
QY 1366 VQVQSLDTVTQTAEAVEKVIETVNSE-----TGSEPECCGAHLPAEKS 1411
Db 13890 ITTIEEBE---TPPOELVQEEIEEIEVEKPAPEOPTDTTFATKSEKPYTEELPEBOV 13946
QY 1412 SATGCHMTLOHAEDTVPLG---PESQAESIPITVPAPSTLHPDQGEISASQRESEB 1468
Db 13947 -----TIQKKKKKAPVEEVEPEAPFLVKRKTVOEVT---BEAKITSKKPVKEE 13995
QY 1469 EDKPDAPGPDAGKESTAIEKVLKAEPIIEEL-----ESKSNKIVLVNITQVAVQPAR 1520
Db 13996 E-----AAELKVTITEEIPTEPEVOEIIIEIEIEEKEPAEYVIEKESQPAVED 14047
QY 1521 TEST-----APETHAVDSO--TQVPACRLDSREPNRCWTQKMDAKM----- 1558
Db 14048 KEVSLPKKKPKAPVIEEPEAEITLKPKVSEEVQBEAKIVKPKKIDEVAVADELTQKV 14107
QY 1559 -KHPVQPREDLQVTLVLEAMQPRKCLPR-----LQLKAP 1593
Db 14108 EEEVVEPEPIVEEVEIEFEIKKKPKPEPEDIVDAIVLTKK 14150

RESULT 13
Q917U4 PRELIMINARY; PRT; 6815 AA.
AC Q917U4;
DT 01-MAR-2001 (TEMBLrel. 16, Created)
DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)
DE 01-MAR-2002 (TEMBLrel. 20, Last annotation update)
DE CG18242 protein.
CN CG18242.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OK NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;

```

```

RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amaratides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.P.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abtill J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Bencos P.V., Bereman B.P., Bhandal D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Brotter P.,
RA Butris J.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Dudin K.J., Evangelista C.C., Ferriz C., Fertiera S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laeok P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Mekulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskeith D.R., Paclet J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reiter K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Morley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195 (2000).
CC -!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
DR EMBL; AE003473; AAC22226.1; -.
DR HSSP; P56276; ITLK.
DR FlyBase; FBgn0035301; CG18242.
DR InterPro; IPR003862; FN1IT_repeat.
DR InterPro; IPR003961; FN1IT.
DR InterPro; IPR003598; IG_C2.
DR InterPro; IPR003600; IG_1like.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR001452; SH3.
DR Pfam; PF00041; fn3; 5.
DR Pfam; PF00047; ig; 11.
DR Pfam; PF00018; SH3; 1.
DR PRINTS; PR00014; FNTYPE111.
DR SMART; SM00060; FN3; 2.
DR SMART; SM00408; IG_C2; 5.
DR SMART; SM00410; IG_1like; 6.
DR SMART; SM00326; SH3; 1.
DR PROSITE; PS50002; SH3; 1.
KW Immunoglobulin domain; Repeat; SH3 domain.
SQ SEQUENCE 6815 AA; 779559 MW; A4E244001A4EBA01 CRC64;

```

```

Query Match 6.9%; Score 557.5; DB 5; Length 6815;
Best Local Similarity 20.5%; Pred. No. 8.9e-14;
Matches 393; Conservative 304; Mismatches 716; Indels 503; Gaps 80;

```

```

QY 7 TQORSPEDPAG-----SDTP-SELIVSGSGPAEASGAGADPADADPARKLPK 54
Db 2989 TTKRMKLRAPGAGEITIEIIVARDPDEABITIVTEP-----EPVNDKKEPKPK 3039
QY 55 NQGLSSVGVAEQGVH--VQ-----EENQGEVEVDE-----DVQGRE 93
Db 3040 KTR-----KYKKDIDHVIQCLIELTPKTELEKYEKTEFPPIYKDKRPSDIVLDES 3093

```


RT TAC and BAC clones".
 RL DNA Res. 7:217-221(2000).
 DR EMBL; AP002057; BAB03174.1; -
 SQ SEQUENCE 2081 AA; 232851 MW; D36031E1F85EFF29 CRC64;
 Query Match 6.4%; Score 517; DB 10; Length 2081;
 Best Local Similarity 19.8%; Pred. No. 9.4e-13;
 Matches 324; Conservative 293; Mismatches 721; Indels 302; Gaps 55;
 Oy 5 SSTEORSPROPASDTPSELVLSGHPPAAEASGADPADADPATKLPOKNQ-----L 58
 Db 191 ASSESSSTHNNVTGSMNV-ETNGENSESTQEKGDVGSNGSDVSMENLGKNVEDL 249
 Oy 59 SSVNGVAGODNVHVEENOGGEEVVDVCGRESEDEVREKRVEMANSTAVADITK 118
 Db 250 KEGNNVVENGE--TKENNENVENNEKEVEGQES-----IGSAIE- 290
 Oy 119 DQGEETSEIIEOIPASENNVEEWD-ASQANDVGFKVFKVGFVKFKDKN----- 172
 Db 291 KULESKEDYKSEVAKAKNGSSMTENLGEAQNN-----GVS-TIDNEKEVEGG 339
 Oy 173 ---EKSDTVQLTVKKDEGAEASVAGDHPESVETAVGESASKESELKOSTEKQEGT 229
 Db 340 ESIJEDDIEKLESKEDYKSEVAKAKNGSSMTGKLEBAQRNNGVSTNETMNSENKGSQ- 398
 Oy 230 LKOEQGSTPILOAESDQAABEAKDE-----GEEKQEK---EPTKSPESSTPY 276
 Db 399 ---ESTNDKQVNAATTNDEBHKENKEETHENNESVKGENTLENKAGNESMKGENLENKY 455
 Oy 277 NSETTSSFFKFTFHGAGNRKKTSPFKSKEDDLETAKEKEQ-----EAEK 322
 Db 456 GNEELK-----GNAASYEAKTNNESSKEEKSORSNEVYMKETTKGENVNIQES 507
 Oy 323 VDEBEKEKTEPAESEOEPAEDTDARLSADYKEVLEPLDQVGDLEASSEKCAPLATEV 382
 Db 508 IODSTDNLENKEDYKPKVDANESDNGSTKERHQ---EAQVNN-GVSTEDKLDNIGAD 563
 Oy 383 FBEKKAHOEVVAEVAHVSTVEKTEBEQGGGGAEGVVGVTGESLPPREKLAEPQYPOE 442
 Db 564 EOKKNDKSVAVTTNDGDHTEKKEETQGNNGES---VKNNLENKEDKELKDDESVGAK 620
 Oy 443 ABAEELMKSRWCVSGD---HTQLTDLSPBEKTLPKHPE---GIVSEVMLSGERIK 496
 Db 621 TNNETSLSEKREBQTKGHNDJNSIKIVDNKGADSNKEKEVHVGSTDNNESEKEDTK 680
 Oy 497 VOGSPILKLFSSSGLKKLGGKQKRGGGDEPEYOHITHTESPASDEQKGESSASS 556
 Db 681 -----SEVAVKNDGSSSEKGBEKENNKDSMEDKTL--ENKESQDSKODKSVDD 728
 Oy 557 PEEPEETCLEKGPLAPODGEAEBCGTSDCG-KKREGITPMASFKKMTTPKRRV--- 611
 Db 729 KOBEAQI-----YGBESKDOKSVEAKGKKES---KENKTKTNNENRYNKKEE 773
 Oy 612 ----RSESDKEBELEKVSATLSTSDS-----TVSEMDQEKTVGEBQKPEPRRYD 661
 Db 774 NVQGNKKESEKKEKESKAKSVETKDNKLSSTENDEAK---EKSGEDNKEDXE 829
 Oy 662 TSVSWEALICVSGSKRARKASSDDGEPRTLGDSH----- 699
 Db 830 ESDYGVSV-----EAKKKNNGVDITNVGNKEDSKDLKDRSVEVKANKEESMK 878
 Oy 700 -RAEASKDEKACTDAVP--ASTQODQOAGSSPPRPAESPGEVSTWESKRLVTPR 756
 Db 879 KRREEVQRWDKSTKEVRPFANMMDIDVQKSGESVKKYKDEKKEG--NKEENKDDITNS 916
 Oy 757 KSKSKLEKAEBSVQJSTLEPSPREEBWSIKKFIPIPRRKRADQKQEOATVDSG 816
 Db 937 SKQKGDKKKKK---ESKNSNKKKEEDK---KEYNNELKQODNKKETTKYSNSKL 989
 Oy 817 VEINEDPVPVAVPLSEVNAVERREKMEAGNTELPOLGAVVSEELSTIVHTVAV 876
 Db 990 KEENKDNKKEKSEDSASKNREKKEVEKKSKTK-----BEAKKEKKKSKQ----- 1034

Oy 877 IDGTAVTVSVERSPSWISASVTEBLEHTAGBAMPVEEVTEKODIABETPVLTOTLPEG 936
 Db 1035 -DKKKEEKDSEBRK-----SKKEEBSRDLAKKKEETYEK-----YESNHSKKED 1083
 Oy 937 KDAHDMVTSEVDFTEAVTATETSEALRTE-----VTEASGAETTDM 981
 Db 1084 KKEHNDKSMKKEEDKKKKKHBSKSKRKKEEDKKDMKLEBDQNSNKKKEDGNEKKKSGH 1143
 Oy 982 VSAVSQLDSPD--TTEATPVQVEVSGVLDTBEERQTOALQAVDKVEESQVPATQ 1039
 Db 1144 VLVYKKESEKKEKKEKKESEKTEKTEIESSKQKNVEYDKKE-----KSKDQOKKEK 1195
 Oy 1040 TVQRTGSKALEKVEE-----VEEDSEVLASEKEDVM---PKGPOEAGAEHLAQSE 1089
 Db 1196 EYKSEBEKLLKKNNEBRKKQTSVEENKQKQETKEKKNRPKODKNTTQSGGKESMESE 1255
 Oy 1090 TQOATPESLEVEVADYDVHATCOVIRLQOLMEQAVAPESSELTLDSETNGSTPLASD 1149
 Db 1256 SKEA--ENQGSQATTQADSDESKNELMQ--ADSQADSHSDSQADSDSKNELMQAD 1310
 Oy 1150 TADGTQODETIDSQDSKATPAVROSQVTEBEAATAQKEEPSTLPNNVPAQ---BEHGE 1205
 Db 1311 SQATTQRNNEBRKKQTSVAENKQKQETKEE-----KKNPKODKKNVTTQSGGKESMES 1365
 Oy 1206 PGADVLEPTQBELTAAPVLAKTVEQGEV-----DMIDGSKVKEEDQVFVHSGN 1258
 Db 1366 ESKENENQKSQATTQADSDESKNELMQADSQADSHSDSQADSDSKNELMQADSOAT 1425
 Oy 1259 SOKAADVTYDSFVMVAGQOEKESTEVQSLSEBEMETDVEKEKRETKPEQVSEGE-- 1316
 Db 1426 TORNNEBRKKQTSVAENKQKQETKEKKNRPKODKNTTQSGGKESM-BESSEAEENQ 1484
 Oy 1317 ---QETAPHEGTGKPYLTIDMPSSERKALGSLGSPSLPDOKA-GCIEVQVSID 1372
 Db 1485 QKSQATTQGESDESKNELMQADSOADTHANSQ-----DDESKNELMQADSOA 1535
 Oy 1373 TTVTQTAEAVEKI---ETVVISETGESPECVGALLPAKSSAAGCHTLOHADIVPL 1429
 Db 1536 DSQTSDESKNELMQADSQADSOAT--DSDESKNELMQADSOAKG---ESLBNKYK 1589
 Oy 1430 GPESQAESIPILVTPAPSTLHPDQGEISASQRESEEDPPADGPADGKSTAIERY 1489
 Db 1590 GKEDNGDEV-----GKNSKTIYKVGHNHESKDKQKTENNGSKVSTBGSQDSNIVERN 1643
 Oy 1490 LKAPELLELESKNKIYLN 1509
 Db 1644 GKEKDSIKESGDKTVEIN 1663
 RESULT 15
 P91257 PRELIMINARY; PRT; 3484 AA.
 AC P91257;
 DT 01-MAY-1997 (Tremblrel. 03. Created)
 DT 01-OCT-2001 (Tremblrel. 18. Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19. Last annotation update)
 DE Hypothetical 385.7 kDa protein.
 GN F12P3.3.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodertinae; Caenorhabditis.
 CX NCBI_TaxID=6239;
 RN (1)
 RP SEQUENCE FROM N. A.
 RC STRAIN=BRISTOL N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA None;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N. A.
 RC STRAIN=BRISTOL N2;

RA Fulton B., Wohldmann P.;
RT "The sequence of C. elegans cosmid F12F3.";
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RN [3]

RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";

RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; U80022; AAC25885.2; -

DR InterPro; IPR001064; Crystallin.

DR InterPro; IPR003962; FhIII repeat.

DR InterPro; IPR003961; FN_III.

DR InterPro; IPR003599; Ig_c2.

DR InterPro; IPR003598; Ig_c2.

DR InterPro; IPR003600; Ig like.

DR InterPro; IPR003006; Ig_MHC.

DR InterPro; IPR000130; Zn_MTpeptdase.

DR Pfam; PF00041; fn3; 7.

DR Pfam; PF00047; ig; 10.

DR PRINTS; PR00014; FNTYPEIII.

DR SMART; SM00060; FN3; 8.

DR SMART; SM00409; IG; 12.

DR SMART; SM00408; IGC2; 8.

DR SMART; SM00410; IG like; 1.

DR PROSITE; PS00225; CRYSTALLIN BETAGAMMA; UNKNOWN_1.

DR PROSITE; PS00142; ZINC PROTEASE; UNKNOWN_1.

DR Hypothetical protein; Immunoglobulin domain; Repeat.

SW Hypothetical protein; Immunoglobulin domain; Repeat.

KQ SEQUENCE 3484 AA; 385746 MW; 21DD6F6893020F4F CRC64;

Query Match 6.2%; Score 501.5; DB 5; Length 3484;
Best Local Similarity 20.3%; Pred. No. 7e-12;
Matches 333; Conservative 285; Mismatches 560; Indels 461; Gaps 69;

QY 80 QEEVEVDVQRESDVREKDRVEEAAANSTAVEDITDKQOETSIIIEQIPASENNV- 138

DB 269 RDGLIDGGSDR--MEVRHD---EMKWLILKIDCKD--EAEYACQAI-----NVA 316

QY 139 -----EEMVQPAESQANDVGFKVFKFVGFKFTVKKDKN-EKSDTVQLLTVKKDEGEG 190

DB 317 GEAWCFSDVVVHMSERDD-----DKSVDEVDSDTVLEEKDDGD-- 357

QY 191 AEASVAGDHOEFSV-ETAVGESASKESELKQSTEQEGLTKQ--EQSSTEIPLQAESDQ 247

DB 358 -----DKSKPKTKKKIKKKEPTPESEQVTAAPPEQQKISEVDVQSVAEVEGAKKP 409

QY 248 AAE-----BAKDEGEKQEKPTKSPESPSS---PVNSETTSFKFFTHGAWGRKKT 299

DB 410 DAEPDLSKAKDKSKSDPEAPESTEETKPTNDKTS-----KKK 454

QY 300 SFKSKEDDLLETAERKEQAEKVDDEBEKTEPASEQEPADTDQARLSADYKVELP 359

DB 455 AEKTVKPKKEVT--GRPLEAKPVEDKDASQPS--SKESPPTDGKKKQIPKALFIP 511

QY 360 --LEDQVGD-----LEASSEK---CAPLATEVFDKMAHOEWAHVH 398

DB 512 DEISSREFGDPSTWHSENNITTTIRREGSADAKTPLVEPLASV-SMKVESAKE---KAE 567

QY 399 VSTVETKEEQQGGEAGGVVGTGESLPPEKLAEPQ-EVQEAEPAPAEELKMSRECV 457

DB 568 FSKRRSTPDDKSRKKEG-----LPPAKKSKKDEVTAEKQSTALIESKKEV 617

QY 458 SGGDHTQLDLSPEKT-----LPHPEG-----IVSEVEMLSQERIIVQSGPLKLF 506

DB 618 ---DEKISQQSDKNKSEVGVPEKAAGETYKQVSEIEV-----PKKTI 663

QY 507 SSGGLKLSGKKQKRGGGDEEPGYQHHTSPESADEQKGESSASSPPEETTC 566

DB 664 KKKTEKSDSSISQSNVLKPADDDKSDVDTKSKKTTEDQTKVATDSKLEAADT-- 721

QY 567 EKGPLEAPQGEAEGETTSGEKKREGITPWASFQKQVTPKVRVRPSDESKDEELKVK 626

DB 722 KQIETETVDDKSKKVL---KKTEKSDSFIQKSETPPVPEPTPAESQAQIAEVNK 778

QY 627 SATLSSTDSTVS-EMQDEVKTVGEEQKPEPPKRRVDTSVSWREALICVGSKKKARKASS 685

DB 779 AKQKEVDONLKREAEVAAKIADEKLIKTEAEANIKKTAEEA-----AKQKEK 828

QY 686 DDE-----GGPRTILGGDSHRAEEASKKEAGTGDVAPASTQODQAOQSSSPPEAGSPS 738

DB 829 DEQLKLETEVSVKKSAAEKLLEBKQAIKKAABADAVKKQKELNEKNKLEAKKSAADKL 888

QY 739 EGEVSTWESFK---RLVTPRKKSLSKLBK-----AEDSSVQLSLSTEIPSPSEESVSI 790

DB 889 KLEESAASAKSVSESVKGEBEKTKAGEKTQVSESEPTSKKTIIDTKDVGATEPADETP 948

QY 791 KKFIPGRKKRADGQEQATVEDSGPV--EINEDDNNVPAVPLSEYNAV-EREKMEAOQ 847

DB 949 KKKIIKKTEKSDSSISQKSATDSKVSQKEQDEPTKPAV---SETQMTVEADSKKKQK 1005

QY 848 NTELPQLLGAIVVSEELSXTLVHTSVAVIDGTRAVTSVEERSPSWISASVTPLEHTAG 907

DB 1006 ETD-----EKLK-----LDAAIAAKTQEAD-----EKS 1031

QY 908 EAMPPVEEVEKTDIIAETPVLTTLPPEGKDAHDMVTSEVDFTSEAVTATETSEALRTE 967

DB 1032 DAQEKIKVSEDDAARKE-----KELND-----KLESEIATKASADKLE 1075

QY 968 EVTEASGAETTDVMSAVSGLTDSPTTBEATPVQEVESGLDTE-----EER 1016

DB 1076 EQAQAKKAAE---VEAAKKQKEKDBQLK-----LQTEAASKKAAAKLELEK 1119

QY 1017 QTAIIQIQAQVAD-----KVVEESQVPAQTQVQ-----RT 1044

DB 1120 QAQIKKAAGADAVKKQKELDKNLEANKKSAAGKLKIEESAASAKQTVEEQAKLDAOT 1179

QY 1045 GSKALEKVEVEEDSEVLASEKEKQVMPKGPVQ-----EACAHLAQSSEGTQATPES 1097

DB 1180 KATAEKQTKLEKDEKSTKESSEKTVDEKPKKKVKKKTEKSDSSISQKSETSKTVVES 1239

QY 1098 LEVPEVTADVHVATCQVILQQLMEQAVAPESSET--LTDS-----ETNGSTPLADS 1150

DB 1240 -----AGPSESETQKVADAAARKQKTEKQKLEAEIT 1271

QY 1151 ADGTQOQDETIDSQSKATAAAROSQVTEBEAATAQKEEFTLNNVPAQEBHEGEPGRDV 1210

DB 1272 AKKSADEK-----SKLEAESLKKAAEVAEAAKKQKEKDEQLKLDT---EASAKAAAEK 1322

QY 1211 LEPTQOELTAAAVPVLA-----KTEVGQGEVDWLDGKVKBEQEVFVHSGPNSQK 1261

DB 1323 LELEKSHIKKAAAEVDVAVKKQKELEBKQLESEAAATKKADAELKLEEQ-----KKK 1374

QY 1262 AADVTYDSEVMGVAGQKEKESTEVEQSLSLEEEMETDVKEKRETKPEQVSEGEQETA 1321

DB 1375 AAEIA---LIEIQKEQEKLAQB-QSRLEDEAKKSAEKQKLESETKSKQTEE-----A 1422

QY 1322 PEHEGYKGPVLTLDMPSSERKALGSLGGSPSLPDQDKAGCTIEVQVQSLDTTWTQ--- 1377

DB 1423 PKE-----SVDEKPKKKVKKKTEKSDSSISQSKS 1453

QY 1378 ---TAAVE-----KVIEVTVISETGESPECVGAHLLPAEKSSATGGHWTLOHAEDTVP 1428

DB 1454 AKSTVDAAEETLESDFNLVEKKTQVKEQSPD-----ESTSATIKRDPAPQKTEISEIK 1504

QY 1429 LGPESQAESIPITVTPAPESTLHPDLQGISASQERSSEEDKPDPA-GPDA----- 1478

DB 1505 QDDGDEKKTITDGGPKPKPDS-----EATPKRVVKKTKQSDVASDASLADSVKLS 1557

QY 1479 DGKESIAIEKVLK-----APPEILELES-KSNKIVLNVIOITAVDOFA 1519

DB 1558 DVVEEKPKKKVKKKTEKSDSIVSETSSVDTTIKPESVEIPTKAEQWILH-----NRFS 1611

QY 1520 RIETAPET---HAYDSQTQ 1535

DB 1612 -TDSAVESEFPKNAHKDDTE 1629

Fri Dec 13 15:34:18 2002

us-09-902-432-4.rpt

Page 18

Search completed: December 13, 2002, 00:16:01
Job time : 166 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: December 13, 2002, 00:17:16 ; Search time 8528 Seconds
(without alignments)
5446.538 Million cell updates/sec

Title: US-09-902-432-4
Perfect score: 8073
Sequence: 1 MGAGSSTEORSPEQAGSDT.....AWAQRKCLPRLQKAPVSK 1596

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool/US09902432/runat_12122002_141451_25800/app_query.fasta_1.1735
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0_1 -LOOPECL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORW=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09902432 @CGN_1_1_6311 @runat_12122002_141451_25800 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_M3AP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.in.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

RESULT 1

ALIGNMENTS

Result No.	Score	Query Match	Length	DB ID	Description
1	8073	100.0	6160	10	U23146 Rattus norv
2	8030	99.5	5236	10	U41453 Rattus norv
3	7461	92.4	178954	2	AC115647 Rattus no
4	6663	82.5	6195	10	AB020886 Mus muscu
5	6663	82.5	6197	10	AF326228 Mus muscu
6	6255	77.5	5902	10	AF326230 Mus muscu
7	6005	74.4	5134	6	AR071282 Sequence
8	4340.5	53.8	6614	6	AX405857 Sequence
9	4300.5	53.3	6608	6	AX329766 Sequence
10	4300.5	53.3	6608	6	AX410793 Sequence
11	4300.5	53.3	6608	9	HSU81607 Sequence
12	4281	53.0	6605	6	AR002603 Sequence
13	4231.5	52.4	6297	9	AF001504 Homo sapi
14	4097	50.7	104939	9	AL590413 Human DNA
15	4097	50.7	171158	2	AC023201 Homo sapi
16	4090	50.7	6287	9	AB003476 Homo sapi
17	1805	22.4	8439	5	AF308810 Xenopus l
18	1585.5	19.6	102822	2	AC097731 Rattus n
19	1067	13.2	1027	9	BC000188 Homo sapi
20	1067	13.2	1027	9	BC022814 Homo sapi
21	807	8.0	160827	3	AC099025 Drosophil
22	639.5	7.9	162921	3	AC007853 Drosophil
23	639.5	7.9	181132	3	AC008206 Drosophil
24	639.5	7.9	227398	3	AE003750 Drosophil
25	635	7.9	16377	3	AF202180 Plasmodiu
26	634	7.9	19362	3	AC007579 Drosophil
27	634	7.9	260367	3	AE003808 Drosophil
28	629	7.8	113474	2	AC013932 Drosophil
29	612.5	7.6	134855	2	AP004077 Oryza sat
30	601.5	7.5	32009	3	CET06D8 Caenorhabdi
31	589.5	7.3	53424	3	AB055861 Procamb
32	583.5	7.2	75313	2	AC017682 Drosophil
33	582	7.2	144056	3	AC091222 Drosophil
34	582	7.2	194006	3	AC010063 Drosophil
35	577.5	7.2	39749	3	DMC49E4 Drosophil
36	576.5	7.1	179145	9	AC009785 Homo sapi
37	576.5	7.1	181360	3	AC104141 Drosophil
38	565.5	7.0	302855	3	AE003420 Drosophil
39	562.5	6.9	68727	2	AC004516 Drosophil
40	557	6.8	23554	2	AC004426 Drosophil
41	547.5	6.7	304106	3	AC024760 Caenorhab
42	537	6.7	304106	3	AE003473 Drosophil
43	536	6.6	70398	3	AE003473 Drosophil
44	533	6.6	125398	3	DM2271740 Drosophil
45	517	6.4	73391	8	AP004150 Oryza sat
					AP002057 Arabidops

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

QY 121 GlnGluThrSerGluIleIleGluGlnIleProAlaSerGluAenAenValGluGlu 140
Db 393 CAGAGAGACATCAGAAATAATTGAACAGATCCCTGCTTCAGAAACAATGTGGAAGAA 452
QY 141 MetValGlnProAlaGluSerGlnAlaAenAspValGlyPheLysLysValPheLysPhe 160
Db 453 ATGGTACAGCCTCGTCTGAGTCCCAAGGCTAATGATGTTGGCTTCAAGAAAGTATTAAATTT 512
QY 161 ValGlyPheLysPheThrValLysLysAspLysAenGlnLysSerAspThrValGlnLeu 180
Db 513 GTTGGTGTAAATTCACGGTGAAGAAGGATAAAAATGAAAAGTCAGATACTGTCCAACTA 572
QY 181 LeuThrValLysLysAspGluGlyGluGlyAlaGluAlaSerValGlyAlaGlyAspHis 200
Db 573 CTCACGTCTCAAGAAGATGAAGCGCAAGGGGAGAGCCCTCTGTGGAGCTGGAGACCAC 632
QY 201 GlnGluProSerValGluThrAlaValGlyGluSerAlaSerLysGluSerGluLeuLys 220
Db 633 CAGAGCCCAAGTGTGGAGACTGCCGTGGAGAGTCAAGATCCCAAGAAAGTGAAGCTGAAG 692
QY 221 GlnSerThrGlnLysGlnGluGlyThrLeuLysGlnGlnSerSerThrGluIlePro 240
Db 693 CAATCCACAGAGAAGCAAGAAGCACCCCTGAAGCAAGACAGAGCAGCACAGAAATCCCC 752
QY 241 LeuGlnAlaGluSerAspGlnAlaAlaGluGluAlaLysAspGluGlyGluLys 260
Db 753 CTTCAAGCCGAATCTGATCAAGCGGCTGAGAAAGACCAAGATGAAGAGAGAAAAA 812
QY 261 GlnGluLysGluProThrLysSerProGluSerProSerSerProValAenSerGluThr 280
Db 813 CAAGAGAAAGAGCCCAACCACTCCCNAGATCCCGAGCAGCCAGTCAACAGTGAAGACA 872
QY 281 ThrSerSerPheLysLysPhePheThrHisGlyTrpAlaGlyTrpArgLysLysThrSer 300
Db 873 ACATCTCTCTTCAAGAAGTTCTTCACTCACGGTTGGCGCGCTGGCGCAAGAAGACCGC 932
QY 301 PheLysLysSerLysGluAspAspLeuGluThrAlaGluLysArgLysGluGlnGluAla 320
Db 933 TTCAGAAATCAAAAGAGGATGATCTGGAACCTGCGAGAAAGAAAGAGCAAGAGGCA 992
QY 321 GluLysValAspGluGluGluLysGluLysThrGluProAlaSerGluGluGlnGluPro 340
Db 993 GAAAAGTAGACGAGGAGAAAGAAAGAAAGACAGAGCCCTCGAGAGGAGCAGAGCCG 1052
QY 341 AlaGluAspThrAspGlnAlaArgLeuSerAlaAspThrGluLysValGluLeuProLeu 360
Db 1053 GCAGAAACACACAGCAGCCAGCCAGTTGTCTCAGAGACTACAGAGAGGTGGAGCTGCCTTG 1112
QY 361 GluAspGlnValGlyAspLeuGluAlaSerSerGluLysCysAlaProLeuAlaThr 380
Db 1113 GAAGACCAAGTTGGTCACTGGAGGATCGTCAAGAGAGAGTGTCTCTCTTGGCAACG 1172
QY 381 GluValPheAspGluLysMetGluAlaHisGlnGluValAlaGluValHisValSer 400
Db 1173 GAAGTGTTCATGAGAAGATGGAAGCCCAAGAAAGTTGTTCAGAGGTCACAGTGAAGC 1232
QY 401 ThrValGluLysThrGluGluGlnGlnGlyGlyGlyGluAlaGluGlyGlyValVal 420
Db 1233 ACCGTGGAGAAGACAGAGGAGGAGCAGGAGGAGGAGGAGGCTGAAGGGGGCGTGGTG 1292
QY 421 ValGluGlyThrGlyGluSerLeuProProGluLysLeuAlaGluProGlnGluValPro 440
Db 1293 GTAGAAGAAACAGAGAGAAATCTTGGCCCTCGAGAAACTGGCTGAGCCCAAGAGGTCCCC 1352
QY 441 GlnGluAlaGluProAlaGluGluLeuMetLysSerArgGluMetCysValSerGlyGly 460
Db 1353 CAGAACTGAGCTGCTGAGGAGCTGATGAAGAGCAGAGAGTGTGTCTCTGGAGGA 1412
QY 461 AspHisThrGlnLeuThrAspLeuSerProGluGluLysThrLeuProLysHisProGlu 480
Db 1413 GACCACACTCACTGACAGACCTAAGTCTCTGAAGAGAGAGCGTGCACCAACACCCAGAA 1472
QY 481 GlyIleValSerGluValGluMetLeuSerSerGlnGluArgIleLysValGlnGlySer 500

Db 1473 GGCAITTTGTCAGTGAAGTGGAGATGCTGTCTCTCAGAAAAGAAATCAAGGTACAGGGAAGT 1532
QY 501 ProLeuLysLysLeuPheSerSerSerGlyLeuLysLysLeuSerGlyLysLysGlnLys 520
Db 1533 CCCTTCAAGAAACTCTTTCAGTAGTCTCAGGCTTAAAGAAAGCTGTCTGGGAAGACAGAA 1592
QY 521 GlyLysArgGlyGlyGlyAspGluGluProGlyGluThrGlnHisIleHisThrGlu 540
Db 1593 GGGAAACGAGGAGGTGGGGAGACGAAAGAGCCCTGGAGAAATACCAACACATTTCAACCGAA 1652
QY 541 SerProGluSerAlaAspGluGlnLysGlyLysSerAlaSerSerProGluGluPro 560
Db 1653 TCCCCAGAGAGTCTGATGAGCAAGAGGAGAGAGTCTTCGCTGCTCCCCCGAGAGACCT 1712
QY 561 GluGluThrThrCysLeuGluLysGlyProLeuGluAlaProGlnAspGlyGluAlaGlu 580
Db 1713 GAGGACACCACTGTCTGGAGAAAGGCGCTGGAAGACCCCGAGGATGGGAAGCTGAG 1772
QY 581 GluGlyThrThrSerAspGlyGluLysLysArgGluGlyIleThrProTrpAlaSerPhe 600
Db 1773 GAAGGAACCTACTTCGATGGAGAAAGAGAGAGAGGATCACTCCCTGGGCATCCTTC 1832
QY 601 LysLysMetValThrProLysLysArgValArgProSerGluSerAspLysGluGlu 620
Db 1833 AAAAGATGGTGACACCCCAAGAAACGGGTCCGAAAGACCTTCTGAGAGTGACAAGAGGAA 1892
QY 621 GluLeuGluLysValLysSerAlaThrLeuSerSerThrAspSerThrValSerGluMet 640
Db 1893 GAGCTGGAGAGGTCAAGAGCGCCACTTGTCTCTCACTGATAGCACAGTGTCAAGAAATG 1952
QY 641 GlnAspGluValLysThrValGlyGluGluGlnLysProGluGluProLysArgArgVal 660
Db 1953 CAAGATGAAGTCAAAACTGTGTGTAGGAAACAAAAGCCAGAGAAACCAAGACGTAGGGTG 2012
QY 661 AspThrSerValSerTrpGluAlaLeuIleCysValGlySerSerLysLysArgAlaArg 680
Db 2013 GATACTTCAGTGTCTTGGGAAGCAGCTGATTTGTGTGGATCATCCCAAGAAAGAGCAAGG 2072
QY 681 LysAlaSerSerSerAspAspGluGlyGlyProArgThrLeuGlyGlyAspSerHisArg 700
Db 2073 AAGGCATCTCTTCAGATGATGAGGAGGGCCCAAGGACACTGGGAGGAGACACTCAGAA 2132
QY 701 AlaGluGluAlaSerLysAspLysGluAlaGlyThrAspAlaValProAlaSerThrGln 720
Db 2133 GCAGAGGAGGCGCAGCAAGAAAGACCAAGAAAGCCGAAACAGACGCTGTCTCTGCAGCACCCAG 2192
QY 721 GluGlnAspGlnAlaGlnGlySerSerSerProGluProAlaGlySerProSerGluGly 740
Db 2193 GAGCAGGACCAAGCGCAAGGAAGTTCTTCCCGAGCCAGCGGGAAGCCCTTCCGAAGGG 2252
QY 741 GluGlyValSerThrTrpGluSerPheLysArgLeuValThrProArgLysLysSerLys 760
Db 2253 GAAGTGTCTCCACTTGGGAGTCATTTAAAGATTAGTCACTCCAAGAAAAAATATCCAAG 2312
QY 761 SerLysLeuGluGluLysAlaGluAspSerSerValGluGlnLeuSerThrGluIleGlu 780
Db 2313 TCAAAACTGGAAGAGAAAGCCGAAAGACTCTAGTGTAGAGCAGTTGTCCACTGAGATCGAA 2372
QY 781 ProSerArgGluGluSerTrpValSerIleLysLysPheIleProArgArgLysLys 800
Db 2373 CCAGGTAGAGAGAAATCTTGGGTTTCCATTAAAGAAATTCATCCCGCAGCGCGGAAGAA 2432
QY 801 ArgAlaAspGlyLysGlnGluAlaThrValGluAspSerGlyProValGluIleAsn 820
Db 2433 AGGGCAGACGGGAAGCAAGAACCAAGCCACTGTGAAGACTCAGGGCCAGTGGAGATTAAT 2492
QY 821 GluAspAspProAsnValProAlaValProLeuSerGluTrpAsnAlaValGluArg 840
Db 2493 GAGGACGACCCCTAATGTCCCGAGCGTGTGCTCTCTGAGTATAATGTCAGTGGAGAGG 2552
QY 841 GluLysMetGluAlaGlnGlyAsnThrGluLeuProGlnLeuGlyAlaValThrVal 860

Db 2553 GAGAGATGGAAGCCAGGGGAATACGAGCTGCCACAGCTGCTGGGGGCTGTGTAACGTG 2612
 QY SerGluGluLeuSerIleThrLeuValHisThrValSerValAlaValIleAspGlyThr 880
 Db 2613 TCCGAGAGAGCTCAGTAAGACTCTGTGCTCCACTGTGAGTGTGCGAGTCAATTAGTGGAGCC 2672
 QY ArgAlaValIleThrSerValGluGluIleArgSerProSerThrIleSerIleAspValIleArg 900
 Db 2673 AAGGAGATGACCAAGTGTCCAAAGCGGTCTCTGCTGATATTCGGCTTCGTAACAGAA 2732
 QY ProLeuGluHisIleThrAlaGlyGluIleMetProProValGluGluValIleArgIleAsp 920
 Db 2733 CCTCTTGAACACACAGCGGGAGAGGCCATGCCACTGTGGAAGAGTCACTGAAAAAGAC 2792
 QY IleIleAlaGluGluThrProValLeuThrGlnThrIleuProGluGlyIleAspAlaHis 940
 Db 2793 ATCATGTGAGAGAAACTCTGTGCTCACCCAGACCTTACAGAGGGTAAAGATGCCCAT 2852
 QY AspAspMetValIleThrSerIleValAspPheThrSerGluAlaValIleThrAlaThrGlu 960
 Db 2853 GACGACATGGTACCAAGTAAAGTGAATTCACCTCAGAAAGTGTGACAGCCACAGAGACC 2912
 QY SerGluAlaLeuArgThrGluGluValIleThrGluAlaSerGlyValAlaGluGluThrThrAsp 980
 Db 2913 TCAGAGGCTCTCCGTAACAGAGTACCGAAGCATCGGGGGCCGAAAGACCAACAGAC 2972
 QY MetValSerAlaValSerGluLeuThrAspSerProAspThrThrGluGluAlaThrPro 1000
 Db 2973 ATGGTGTCCGCGATTTCCAGCTGACTGACTCCCCAGACACCAAGAGAAAGCAACCCCA 3032
 QY ValGluGluValGluSerGlyValLeuAspThrGluGluGluArgGlnThrGlnAla 1020
 Db 3033 GTTCAGAGAGTGAAGTGTGTGTGTAATACAGAAAGAGAGGCGCAAGAGCGAGCC 3092
 QY IleLeuGlnAlaValAlaAspIleValIleGluGluSerGlnValProAlaThrGlnThr 1040
 Db 3093 ATCTTCAAGCCGCTTGCACACAGAGTGAAGAGAGTCCAGAGTGTGCTGCAACCCAGACT 3152
 QY ValGlnArgThrGlySerIleValAlaLeuGluIleValGluGluValIleGluGluAspSerGlu 1060
 Db 3153 GTGCAGAGAAAGCGGGTCAAAAGCACTGGAAGAGTGAAGAGTGAAGAGGAGGACTCCGAA 3212
 QY ValLeuAlaSerGluIleArgIleAspValMetProIleGlyProValGlnGlnAlaGly 1080
 Db 3213 GTGCTGGCTTCGAGAAAGAGAGAGAGGATGCTGCAAGAGACCGCTGAGAGGCTGA 3272
 QY AlaGluHisLeuAlaGlnIleSerGluThrGlyGlnAlaThrProGluSerLeuGlnVal 1100
 Db 3273 GCTGAGCATCTTGCAAGAGGCTTGAGACTGAGACAGGCTTCCACAGAGCCCTTGAAGTT 3332
 QY ProGluValIleThrAlaAspValAspHisValAlaThrCysGlnValIleIleLeuGlnGln 1120
 Db 3333 CCTGAAGTACAGGCAATGATGACATGTCGACAGCTGCAAGGTATCAAAAGCTCCAGCAG 3392
 QY LeuMetGluGlnAlaValAlaProGluIleSerGluThrGlyGlnAlaThrProGluSerLeu 1140
 Db 3393 CTGATGAACAGAGCGCTGCGCTCTGAGTATCGAAACCTTGACAGACAGTGAACAAT 3452
 QY GlySerThrProLeuAlaAspSerAspThrAlaAspGlyThrGlnGlnAspGluThrIle 1160
 Db 3453 GGAAGACCTCCCTTACAGATTCAGACCTGCAAGATGGGACACAGCAAGATGAACATTT 3512
 QY AspSerGlnAspSerIleValIleThrAlaAlaValAlaGlnSerGlnValIleThrGluGlu 1180
 Db 3513 GACAGCCAGACAGTAAGCACTGACAGCTGACAGCAAGTCAAGAGTCAAGAAAGAGAG 3572
 QY AlaAlaThrAlaGluIleArgGluGluProSerThrLeuProAsnAsnValProAlaGlnGlu 1200
 Db 3573 GCGGCTACTGCTCAGAAAGAGAGCGCTTGCACCTACTAATATGTCTTCCAGGCCAGGAA 3632
 QY GlnHisGlyGluGluProGluArgAspValIleGluGluProThrGlnGlnIleuThrAla 1220
 Db 3633 GAACATGGGGAAGAACAGGAGAGATGTTCTTGAACCTACACAGCAAGAGCTTACTGCT 3692

QY AlaAlaValProValLeuAlaIleThrGluValGluGlnGluGluValAspThrIleu 1240
 Db 3693 GCAGCCGTGCCCCGTTCGGCAAAAGCTGAGGTGGTCAAGAGGGGTGAGTGAATGCTGTG 3752
 QY AspGlyIleValValIleGluGluGlnIleValPheValHisSerGlyProAsnSerGln 1260
 Db 3753 GATGAGAAAAAGTCAAAAGAAAGACAGAGGTGTTTGTACCTCTGGACCCCAAGCTGA 3812
 QY LysAlaAlaAspValIleThrTyraAspSerGluValMetGlyValAlaGlyCysGlnIle 1280
 Db 3813 AAGCTCTGATGTGACATATGACAGTGAATGAGAGTGGAGCGCGGTGTACAGGAAAAAG 3872
 QY GlnSerThrGluValGlnSerLeuSerLeuGluGluGluValMetGluThrAspValGlu 1300
 Db 3873 GAGAGTACTGAAGTGAAGTCAAGTCTTGAAGGAGGAGAGATGAACCTGACCTTAA 3932
 QY LysGluIleValGluThrIleValProGluGlnValSerGluGluGluGluGlnIleThrAla 1320
 Db 3933 AAGGAGAAAAAGGAGACAAAGCCAGAGCAAGTGAAGAAAGTGAAGAGAGAGAAACAGCC 3992
 QY AlaProGluHisGluGlyThrTyrlYleProValLeuThrLeuAspMetProSerSer 1340
 Db 3993 GCTCTGAGCATGAAGAAACCTACGGAAGCCAGCTGACACTGACATGACATGCCAGCTCA 4052
 QY GlnArgGlyIleValLeuGlySerLeuGlyIleSerProSerLeuProAspGlnAspIle 1360
 Db 4053 GAGAGGGGAGAGCACTGGAGAGCTTGGAGAGAGCCCTTCTTCCACAGACCAACAA 4112
 QY AlaGlyCysIleGluValGlnValGlnSerLeuAspThrThrValIleThrAlaGlu 1380
 Db 4113 GCAGGTTGATAGAGTTCMAAGTTCMAAGCTTCGACACACAGTCACTCAAAACAGCAGAA 4172
 QY AlaValGluIleValIleGluThrValIleSerGluThrGlyIleSerProGluCys 1400
 Db 4173 GCTGTGAAAAAGTCAATGAACCGGTTGATTCAGAGACAGTGGAAGTCCAGAGTGT 4232
 QY ValGlyAlaHisLeuLeuProAlaGluIleSerSerAlaThrGlyGlyHisIleThrIleu 1420
 Db 4233 GTAGGTGACACTTATTCACAGCTGAGAAAGTCTCTTGAACGGGTGACCTGAGCTCT 4292
 QY GlnHisAlaGluAspThrValProLeuGlyProGluSerGlnAlaGluSerIleProIle 1440
 Db 4293 CAGCATGCAAGAGACACGCTACCTCGGGGCTGTGAGTCTCAGGCGAAATCCATCCCAATC 4352
 QY IleValIleThrProAlaProGluSerThrLeuHisProAspLeuGlnGlyIleSerAla 1460
 Db 4353 ATAGTAACTCCGCTCTGAAAGACCTTACATCTGACCTTACAGAGAGAAATTAAGCGCA 4412
 QY SerGlnArgIleArgSerGluGluGluAspIleProAspAlaGlyProAspAlaAspGly 1480
 Db 4413 TCCAGAGAGAGCGATTCAGAGAGAGAGACCAAGATGCTGCTGATGCTTCAAGCGC 4472
 QY LysGluSerThrAlaIleGluIleValLeuIleValIleGluIleValIleGluIleuGlu 1500
 Db 4473 AAGGAGATGACGAAATGAAAGAAAGTCTCAAGGCTGAACCTGAAATCTTGAACCTTGA 4532
 QY SerIleSerAsnIleValLeuAsnValIleGlnIleAlaValAspGlnPheAlaArg 1520
 Db 4533 AGTAAGAGACAAAGATTTGCTGAACGTCAATTCAGACAGCGGTGACCAAGTTCCACAGT 4592
 QY ThrGluThrAlaProGluIleThrHisAlaTyraAspSerGlnThrGlnValProAlaCysArg 1540
 Db 4593 ACAGAAACAGCCCCGAAATCATGCTTATGATTCACAGACCCAGAGTTCTGCAATGACAG 4652
 QY LeuAspSerArgGluProAsnArgCysTrpThrIleMetIleAspAlaIleMetIleHis 1560
 Db 4653 CTTGACAGAGGAGGCCAACAAGATGCTGGACAAAGATGAAGATGCCAAGATGAACAC 4712
 QY ProValProGlnProArgIleAspLeuGlnValLeuThrValIleGluValIleThrAlaGln 1580
 Db 4713 CCAAGTCCGCAACCCAGAGAGGACTTGAAGTCTGACCGGTTCTGGAAGCATGGGCTGAG 4772

QY 1581 ProArgLysCysLeuProArgLeuGlnLeuLysAlaProValSerLys 1596
|||||
Db 4773 CCTCGAATGCTTGCCTGCTGCAAGTGAAGCGCGGTGTCAAAG 4820

RESULT 2

RNU41453 5236 bp mRNA linear ROD 12-JAN-1999
LOCUS Rattus norvegicus PKC binding protein and substrate mRNA, complete
DEFINITION cds.
ACCESSION U41453
VERSION U41453.1 GI:4096819
KEYWORDS
SOURCE Rattus norvegicus.
ORGANISM Rattus norvegicus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE

1 (bases 1 to 5236)
Chapline, C., Mousseau, B., Ramsay, K. and Jaken, S.
Identification of a major protein kinase C binding protein and
substrate in rat embryo fibroblasts: Decreased expression in
transformed cells
J. Biol. Chem. (1995) In press
2 (bases 1 to 5236)
Chapline, C., Mousseau, B., Ramsay, K. and Jaken, S.
Direct Submission
Submitted (28-NOV-1995) Susan Jaken, W Alton Jones Cell Science
Center, Old Barn Road, Lake Placid, NY 12946, USA
Location/Qualifiers

FEATURES

source

CDS

1..5236
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone="72"
33..4823
/note="major PKC binding protein in REF52 cells"
/codon_start=1
/product="PKC binding protein and substrate"
/protein_id="AAB03788.1"
/db_xref="GI:4096820"

/translation="MGAGSTSEQRSPQAGSDTPSELVLSHGHPAAEASGAAGDPAD
ADPAPKLPKNGSSVGVAEQGVHVEQNEQEEVDEVDVQSESDVREKDR
VREMANSTAVEDITKDGQETSEIIQIPASENNVEEMVQPAEQANDVGPKYKPF
VGFKFTVKDKQESDVTQVLLVKKDEGGAASVAGDHQPSVETAVGESKESSE
LKQSTBKGLQKQESSTETPLQAESDQAAEEAKDEGEKEKEPTKSPSPSPV
NSETSPFKFTGHWAGWRKFTSKSKEDDLLETAERKEQEAKEVDEERKEKTEPA
SSEQPAEDTQARLSADKVELPLEDQVGLAESSEKCAPLATEVDFDEKWEAHEM
VVAEYHVSTVEETEEQGGGAEGGVVVEGTGSELPPEKLAEPORVPOEAAPAEELM
KREMCVSGDHTQTLDSPEKTLPHPEGIVSEVEMLSQERIKVQSPKLLFSS
SLKLLSKKKQKRGCGGDEPGEYQHHTSPESADQKGESSASSPEEPEETCL
EXGLPAPQDGAEEGTTSDGKKREGITPWASFQKQVTPKKVRVRPSSEDKKEELEK
VKSATLSDTSTVSEMODVKTVEEQKEPPEPKRVDYTSVWEALICVGSKKRKA
SSDDRGGRPTLGGSHRAEASKDEAGTDVAPVASTQPDQQAQSSSPPEPAGSPSEG
EGSVTWSEFKLVTPRKSKSLKEKAEDSSVEOLSTETEPSRESWVSIKKFIPGR
KRADGQSOATVDSGVPVINEEDPDVAVVPLSEYDAVEREKWEAQNAELPQLLG
AYVSEBELSKTLHVSVAVIDGTRAVTSVEERSQSWISASVTEPLEHTAGAMPVE
EVTEKDIAEETPLVLTLPQEGKADDDMTSEVDFTSVAVTATSETSEALTEEVTEA
SGAETDMVSQVLTSDPTTEATPVQESGVLDTEEBERQTQAILQAVADVKV
SSEQVATQVTSQSKALEKEVEEDSEVLASEKEDVMPKPVQEAAGHLAGQS
ETGAQTPLESYLPVTDVHATCOVIKLOLMEOAVAPESETLTDSETNGSTPLA
DSDTADGTOQDETIDSQDSKATAAVRQSQVTEEEATAQKEPSTLNNVPAQBEHGE
EGRDVLETFQELTAAAPVLAKEVQGEVMDLDEKVEEEDVFVHSGPNSQKA
ADVTYSEYVNGVAGQEESEVQSLSLEGEMETDVEKERETPEQVSEBGEQETA
APEHEGYCKVPLTLDMPSERGKALSGSPSLPDQKAGCIEVQVSLDTVTOT
ASAVEKVIETVVISSETGESPECVGAHLLPAEKSSATGQHWTLQHAEDTVPLGPESQAE
SIPIIVTPAETSLHPDLOGEISASQRESSEBEDKPDAGPDAGKESTAEIKVLKAE
EILELESKNKVLNVIQVAVDQFARTETAPETHAYDSQTQVPMQADSEPNRCWTK
MKVAKMKHPVPOREDLQVLTLEAWAQRKCLPLRLQLKAPVSK"

BASE COUNT 1622 a 1206 c 1583 g 825 t

Alignment Scores:

Pred. No.: 9,42e-279 Length: 5236
Score: 8030.00 Matches: 1589

Percent Similarity: 99.75% Conservative: 3
Best Local Similarity: 99.56% Mismatches: 4
Query Match: 99.47% Indels: 0
DB: 10 Gaps: 0
US-09-902-432-4 (1-1596) x RNU41453 (1-5236)
QY 1 MetGlyAlaGlySerSerThrGluGlnArgSerProGluGlnProAlaGlySerAspThr 20
Db 33 ATGGCGCAGGCGATTCACCGAGCAGCGAGCCCGCAGCAGCGCGCGGAGCGACACG 92
QY 21 ProSerGluLeuValLeuSerGlyHisGlyProAlaAlaGluAlaSerGlyAlaAlaGly 40
Db 93 CCAGGCGAGCTGTGTCTAGTGGCCATGGCCCGAGCTGAAAGCCCTCGGAGCAGCTGGA 152
QY 41 AspProAlaAspAlaAspProAlaThrLysLeuProGlnLysAsnGlyGlnLeuSerSer 60
Db 153 GACCCCGCGCGCGCGACCCCGCCCAAGCTCCACACAGAGAATGCCAGCTGTCTCT 212
QY 61 ValAsnGlyValAlaGluGlnGlyAspValHisValGlnGluAsnGlnGluGlyGln 80
Db 213 GTCAACGGCGGTAGCTGAACAAGGAGATGTCATGTCCAAGAGGAAACCCAGGAGGCG 272
QY 81 GluGluGluValValAspGluAspValGlyGlnArgGluSerGluAspValArgGluLys 100
Db 273 GAGGAAGAAGTCTGTGATGAGGATGTTGGACAGCGAGATGTCAGAAAGATGTGAGAGAAA 332
QY 101 AspArgValGluGluMetAlaAlaAsnSerThrAlaValGluAspIleThrLysAspGly 120
Db 333 GACCGAGTGAAGAAATGGCGGCAACTCCACAGCTGTTGAAGATATCACAAGAGATGG 392
QY 121 GlnGluGluThrSerGluIleLeuGluGlnIleProAlaSerGluAsnValGluGlu 140
Db 393 CAGGAGGAGACATCAGAAATAATTGAACAGATCCCTGCTTCAGAAACAATGTGGAAGAA 452
QY 141 MetValGlnProAlaGluSerGlnAlaAsnAspValGlyPheLysLysValPheLysPhe 160
Db 453 ATGGTACAGCCCTGCTCAGTCCCAGGCTAATGATGTTGGCTTCAAGAAAGATATTAAAT 512
QY 161 ValGlyPheLysPheThrValLysLysAspLysAsnGluLysSerAspThrValGlnLeu 180
Db 513 GTTGGTTTAAATTCACCGTGAAGAGGATTAATTAAGTCAAGTACTGTCTCAACTA 572
QY 181 LeuThrValLysLysAspGluGlyGluGlyAlaGluAlaSerValGlyAlaGlyAspHis 200
Db 573 CTCACTGTCAAGAAGGATGAAGCGAAGGCGAGAACCTCTGTGCGAGCTGGAGACCAC 632
QY 201 GlnGluProSerValGluThrAlaValGlyGluSerAlaSerLysGluSerGluLys 220
Db 633 CAGGAGCCAGTGTGGAGACTCCGTCGAGAGTCAAGCATCCAAAGAAAGTGAAGTGAAG 692
QY 221 GlnSerThrGluLysGlnGluGlyThrLeuLysGlnGluGlnSerThrGluIlePro 240
Db 693 CAATCCACAGAGAGCAGCAAGCACCTCGAAGCAAGAAACAGACGACACAGAAATCCCC 752
QY 241 LeuGlnAlaGluSerAspGlnAlaAlaGluGluGluAlaLysAspGluGlyGluLys 260
Db 753 CTTCAAGCCGAATCTGATCAAGCGCTGAGGAAGAACCAAGATGAAGAGAGAGAAAA 812
QY 261 GlnGluLysGluProThrLysSerProGluSerProSerSerProValAsnSerGluThr 280
Db 813 CAAGAAAGAGGCCCAACCAAGTCCCCCAGAAATCCCCCAGCAGCCAGTCAACAGTGAACA 872
QY 281 ThrSerSerPheLysLysPhePheThrHisGlyThrAlaGlyTrpArgLysThrSer 300
Db 873 ACATCTCTCTCAAGAAAGTCTTCTACTACGGTGGCGCGCTGGCGCAGGAAGACACG 932
QY 301 PheLysLysSerLysGluAspLeuGluThrAlaGluLysArgLysGluGlnGluAla 320
Db 933 TTCAAGAAATCAAAAGAGAGATGATCTGGAAACTGCCGAGAGAGAAAGAGAGAGGCA 992
QY 321 GluLysValAspGluGluGluLysGluLysThrGluProAlaSerGluGluGlnGluPro 340
|||||

Db 993 GAAAACTTAGACGAGAGAAAAAGAGAGCCGCTCGGAGAGAGAGACCG 1052
Qy AAGUAAspThrAspGlnAlaArgLeuSerAlaAspTyrGluLysValGluLeuProLeu 360
Db 1093 GCAGAGAGACACAGACAGGCGGCTTGTTCAGACAGACTACGAGAGAGGTGAGCTGCTTTG 1112
Qy 361 GUAAspGlnValGlyAspLeuGlnAlaSerSerGluGlnLysCysAlaProLeuAlaThr 380
Db 1113 GAAGACCAAGGTGTGACCTGGAGGATGTCAGAGAGAAAGTGTCTCTTTGGCAAG 1172
Qy 381 GUAValPheAspGluLysMetGluAlaHisGlnGluValAlaGluValAlaHisValSer 400
Db 1173 GAAGCTTTGATGAGAGAGTGAAGCCCAAGAAAGTTGTCAGAGAGTCCACGTGAGC 1232
Qy 401 ThrValGluLysThrGluGlnGlnGlnGlyGlyGlyGluAlaGluGlyValVal 420
Db 1293 ACCGTGAGAGAGACAGAGAGAGGAGGAGAGAGAGAGAGGCTGAAGGGGCGTGTG 1292
Qy 421 ValGluGlyThrGlyGluSerLeuProGluLysLeuAlaGluProGlnGlnValPro 440
Db 1293 GTAGAGAGGAACAGAGAAATCTTGGCCCTGAGAAATGGCTGAGGCCAGAGAGTCCCC 1352
Qy 441 GlnGluAlaGluProAlaGluGlnLeuMetLysSerArgGluMetCysValSerGly 460
Db 1353 CAGAGAGCTGAGCCTGCTGAGAGCTGATGAAGACAGAGAGATGTGTCTCTGAGAG 1412
Qy 461 AspHisThrGlnLeuThrAspLeuSerProGluGlnLysThrLeuProLysHisProGlu 480
Db 1413 GACCAACACTCAACTGACACACCTTAAGTCTGAAGAGAACGCTGCCAAACACCCAGAA 1472
Qy 481 GlnLysLeuValSerGluValGluMetLeuSerSerGlnGlnLysValGlnGlnLys 500
Db 1473 GGCATTGTTCAGTGAAGGTGAGATGCTGTCTCTCAGAGAAAGATCAAGGTACAGGGAAAT 1532
Qy 501 ProLeuLysLeuPheSerSerSerSerGlyLeuLysLeuSerGlyLysLysGlnLys 520
Db 1533 CCTTGAAGAACTCTTCTGATGACTCAGGCTTAAAGAACTGTCTGGAGAAAGCAGAA 1592
Qy 521 GlnLysArgGlyGlyGlyLysAspGluGlnProGlyGlnLysIleHisThrGln 540
Db 1593 GGGAAACGAGAGAGTGGGGAGACGAAAGACCTGGAGAAATCAACACATTCACACGAA 1652
Qy 541 SerProGluSerAlaAspGluGlnLysGlyGluSerSerAlaSerSerProGluGlnPro 560
Db 1653 TCCCCAGAGAGTGTGATAGAGAGAGAGAGCTGTGCTGCTCCCGAGGAGCTT 1712
Qy 561 GlnGluThrCysLeuGlnLysGlyProLeuGluAlaProGlnAspGlyGluAlaGln 580
Db 1713 GAGAGAGACCAAGTGTGAGAGAGGGCGCTGGAGACCCCAAGATGGGAGAGCTGAG 1772
Qy 581 GlnGlyThrThrSerAspGlyGlnLysLysArgGlnGlyIleThrProTyrAlaSerPhe 600
Db 1773 GAAGAACTACTTCGATGGAGAGAGAGAGAGAGGATCACTCCCGGGGATCTTTC 1832
Qy 601 LysLysMetValThrProLysLysArgValArgArgProSerGluSerAspLysGlnGlu 620
Db 1833 AAAAAGATGTGACACCCCAAGAAACGGGTCCGAAAGACTTCTGAGAGTGAACAAGGAGAA 1892
Qy 621 GlnLeuGlnLysValLysSerAlaThrLeuSerSerThrAspSerThrValSerGluMet 640
Db 1893 GAGCTGGAAGAGTCAAGAGCGCACCTGTGCTCCACTGATAGCAACAGTTCAGAAATG 1952
Qy 641 GlnAspGlnValLysThrValGlyGlnGlnLysProGlnGlnLysProLysArgArgVal 660
Db 1953 CAAGAGAAAGTCAAACTGTGTGTGAGAAACAACAAGCCAGAGAGACAACAAGCTTAGAG 2012
Qy 661 AspThrSerValSerTyrGlnAlaLeuIleCysValGlySerSerLysArgAlaArg 680
Db 2013 GATACTTCACTGCTTGGAGAGCACTGATTTGTGTGATCATCCAAAGAGAGAGCAAG 2072
Qy 681 LysAlaSerSerSerAspAspGlnGlyLysProArgThrLeuGlnGlyLysSerHisArg 700
Db 2073 AAGGCACTCTCTTCAATGATGAAGAGGGCCAAAGCACTGGAGAGGGGCACTCAACAGA 2132

Qy 701 AlagluGlnAlaSerLysAspLysGlnAlaGlyThrAspAlaValProAlaSerThrGln 720
Db 2133 GCAGAGAGGCGCACCAAGAACAAAGAGCCGGAACAGACGCTGTCTCTCGACAGACCCAG 2192
Qy 721 GlnGlnAspGlnAlaGlnGlySerSerSerProGluProAlaGlnLysSerProSerGlnGly 740
Db 2193 GAGCAGAGCAAGCGCAAGAGAAAGTTCTTCAACCCAGCCAGCGGAGAGCCCTTCCGAAGG 2252
Qy 741 GlnGlyValSerThrTyrGluSerPheLysArgLeuValThrProArgLysLysSerLys 760
Db 2253 GAAGGTGTCTCACTTGGAGACTTAAAGATTAGTCACTCCAAAGAAAAAATTCCAAG 2312
Qy 761 SerLysLeuGlnGlnLysValGluAspSerSerValGlnGlnLeuSerThrGlnIleGln 780
Db 2313 TCMAAATCGAAGAGAAAGCCGAAGACTCTGTGTAGAGCGATTGTCCACTGAGATCGAA 2372
Qy 781 ProSerArgGluGluSerTyrTyrValSerIleLysLysPheIleProGlyArgArgLys 800
Db 2373 CCGAGTACAGAAAGATCTTGGGTTTCCATTAAAGAAATTCATCCCGAGCGCGAGAGAA 2432
Qy 801 ArgAlaAspGlyLysGlnGlnAlaThrValGluAspSerGlyProValGlnIleAsn 820
Db 2433 AGGCGAGTGGGAGAGCAAGAACAGCCACTGTGAGAACTCAGGGCCAGTGAATTAAT 2492
Qy 821 GluAspAspProAsnValProAlaValValProLeuSerGluTyrAsnAlaValGluArg 840
Db 2493 GAGGACGACCTCGATGTCCAGCCGTCGTGCTGTCTGATGATGATGATGATGAGAGG 2552
Qy 841 GlnLysMetGluAlaGlnGlyAsnThrGluLeuProGlnLeuLeuGlnValAlaTyrVal 860
Db 2553 GAGAGATGAGAGCCCAAGGGAAATCGAGAGCTCCCAAGCTGTGGGGCTGTGTACTG 2612
Qy 861 SerGlnGluLeuSerLysThrLeuValHisThrValSerValAlaValIleAspGlyThr 880
Db 2613 TCCAGAGAGCTCAGTAAGACTGTGCTCACACTGTGAGTGTGCGAGTATGATGATGAGAC 2672
Qy 881 ArgAlaValThrSerValGluGlnLysSerProSerTyrIleSerAlaSerValThrGln 900
Db 2673 AGGCGAGTCCACAGTGTGAAGAGCGGTCTCTCTGTGTGAATATCGCTTCGTAAACA 2732
Qy 901 ProLeuGlnHisThrAlaGlyGluAlaMetProProValGlnGluValThrGluLysAsp 920
Db 2733 CCTCTTGAACACACAGCGGAGAGCCATGCCACTGTGAAAGAGTCACTGMAAAAGAC 2792
Qy 921 IleIleAlaGluGlnThrProValLeuThrGlnThrLeuProGlnGlyLysAspAlaHis 940
Db 2793 ATCATTTGCAAGAAAGAACTCTGTGCTCACCCAGAGCTTACCAGAGGTTAAAGATCCAT 2852
Qy 941 AspAspMetValThrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThr 960
Db 2853 GACGACATGTCTCACCGAGTGAATTTCACTCAGAAAGCTGTGACACAGAGACC 2912
Qy 961 SerGluAlaLeuArgThrGlnGluValThrGluAlaSerGlyAlaGlnGluThrThrAsp 980
Db 2913 TCAGAGGCTCTCGTACGAAAGATTAACGAGATGAGGGGCGGAAGAGCCACAGAGC 2972
Qy 981 MetValSerAlaValSerGlnLeuThrAspSerProAspThrThrGluGlnAlaThrPro 1000
Db 2973 ATGTGTCCGACGTTTCCAGCTGACTGATCCCAAGACACACAGAGAGAGCCACCCA 3032
Qy 1001 ValGlnGlnValGluSerGlyValLeuAspThrGlnGlnGlnGlnArgGlnThrGlnAla 1020
Db 3033 GTTCAGAGAGTGAAGAGTGTGTCTAATACAGAAAGAGAGAGCGGACGAGCGAGCC 3092
Qy 1021 IleLeuGlnAlaValAlaAspLysValLysGlnGlnGlnGlnValProAlaThrGlnThr 1040
Db 3093 ATCTTCCAAAGCCGTTCAGACAGAGTGAAGAGAGTCCCAAGTCTGTCAACCCAGACT 3152
Qy 1041 ValGlnArgThrGlySerLysAlaLeuGlnLysValGlnGlnValGlnGlnLysAspSerGln 1060
Db 3153 GTGCAAGAGACGGGTCAAAACCACTGAGAGAGTTCGAGAGAGTTCAGAGAGACTCCGA 3212

QY 1061 ValLeuAlaSerGluLysGluLysAspValMetProLysGlyProValGlnGluAlaGly 1080
 DB 3213 GTGCTGCTTCGGAGAAGAAGAGAGAGCTTATGCGGAAGAGACCCGTCAGGAAGCTGGA 3272
 QY 1081 AlaGluHisLeuAlaGlnGlySerGluThrGlyGlnAlaThrProGluSerLeuGluVal 1100
 DB 3273 GCTGAGCATCTTGACAGGCTCTGAGACTGGACAGAGCTACTCCAGAGAGCCTTGAAGTT 3332
 QY 1101 ProGluValThrAlaAspValAspHisValAlaThrCysGlnValIleLysLeuGlnGln 1120
 DB 3333 CCTGAAGTCACGGCAGATGTAGACCATGTGCCACGTCGCCAGGTATCAAGCTCCAGCAG 3392
 QY 1121 LeuMetGluGlnAlaValAlaProGluSerSerGluThrLeuThrAspSerGluThrAsn 1140
 DB 3393 CTGATGGAACAGGCGGTGGCCCTCGAGTCATCCGAAACCTTTGACAGACAGTGAGACAAAT 3452
 QY 1141 GlySerThrProLeuAlaAspSerAspThrAlaAspGlyThrGlnGlnAspGluThrIle 1160
 DB 3453 GGAAGCACTCCCTTAGCAGATTACAGACTGCAGATGGGACACAGCAAGATGAACCATT 3512
 QY 1161 AspSerGlnAspSerLysAlaThrAlaAlaValArgGlnSerGlnValThrGluGluGlu 1180
 DB 3513 GACAGCCAGGACAGTAAGCCACTGCAAGCTGTCAAGCAGTGCACAGAGTGCACAGAAGAAG 3572
 QY 1181 AlaAlaThrAlaGlnLysGluGluProSerThrLeuProAsnValProAlaGlnGlu 1200
 DB 3573 CGGCTACTCTCAGAAGAGGAGCCCTTCGACACTACTAATAATGTTCCAGGCCCGAGAA 3632
 QY 1201 GluHisGlyGluGluProGlyArgAspValLeuGluProThrGlnGlnGluLeuThrAla 1220
 DB 3633 GAACATGGGAAGAACACAGGAAGAGATGTTCTTGAACTACACAGCAAGAGCTTACTGCT 3692
 QY 1221 AlaAlaValProValLeuAlaLysThrGluValGlyGlnGluGlyGluValAspTrpLeu 1240
 DB 3693 GCAGCCCTGCGCGTCTCGCAAGACTGAGTGGGTCAAGAGGGTGAGGTGACTGCTGTG 3752
 QY 1241 AspGlyGluLysValLysGluGluGlnGluValPheValHisSerGlyProAsnSerGln 1260
 DB 3753 GATGGAGAAAAGTCAAGAAGAACACAGAGGTGTTGTACACTCTGGACCCCAACAGTCAA 3812
 QY 1261 LysAlaAlaAspValThrTyroAspSerGluValMetGlyValAlaGlyCysGlnGluLys 1280
 DB 3813 AAGCTCTGATGTGACATATGACAGTGAAGTGAAGTGGCGGTGTCAGGAAGAAG 3872
 QY 1281 GluSerThrGluValGlnSerLeuSerLeuGluGluGlyGluMetGluThrAspValGlu 1300
 DB 3873 GAGAGTACTGAAGTGCAGAGTCTTAGCTCGAGAGGAGAGATGGAAACTGACGTTGAA 3932
 QY 1301 LysGluLysArgGluThrLysProGluGlnValSerGluGluGlyGluGlnGluThrAla 1320
 DB 3933 AAGGAGAAAAGGGAGACAAAGCCAGACAGCAAGTGAAGAAAGGTGAGCAAGAACAGCC 3992
 QY 1321 AlaProGluHisGluGlyThrTyroLysProValLeuThrLeuAspMetProSerSer 1340
 DB 3993 GCTCTGAGCATGAGGAACTACGGGAAGCCAGTCTCTGACACTTGTGATGCCAGCTCA 4052
 QY 1341 GluArgGlyLysAlaLeuGlySerLeuGlyCysSerProSerLeuProAspGlnAspLys 1360
 DB 4053 GAGAGGGGAAGGACCTGGGAAGCCTTGGAGGAAGCCCTTCTCTCCAGACCAAGACAAA 4112
 QY 1361 AlaGlyCysIleGluValGlnValGlnSerLeuAspThrThrValThrGlnThrAlaGlu 1380
 DB 4113 GCAGGTGTCATAGAGGTTCAAAGTTCAAAGCTCGGACACACACAGTCACTCAACACAGCAGAA 4172
 QY 1381 AlaValGluLysValIleGluThrValValIleSerGluThrGlyGluSerProGluCys 1400
 DB 4173 GCTGTGAAAAGGTCAAGAAACCGTTGTGATTTACAGACAGAGTGAAGTCCAGAGTGT 4232
 QY 1401 ValGlyAlaHisLeuLeuProAlaGluLysSerSerAlaThrGlyGlyHisTrpThrLeu 1420
 DB 4233 GTAGTGCACACTATTATTACAGCTGAGAAGTCTCTCTGCAAGGGTGGCCACTGGACTCTT 4292
 QY 1421 GlnHisAlaGluAspThrValProLeuGlyProGluSerGlnAlaGluSerIleProIle 1440

DB 4293 CAGCATGCAGAGGACACGGTACCCCTGGGGCCTAGTCTCAGGAGAGANTCCATCCCAATC 4352
 QY 1441 IleValThrProAlaProGluSerThrLeuHisProAspLeuGlnGlyGluIleSerAla 1460
 DB 4353 ATAGTAACTCCTGCTCTCTGAAAGCACCCCTACATCTCTGACCTACAAGGAGAAATAAGCGCA 4412
 QY 1461 SerGlnArgGluArgSerGluGluAspLysProAspAlaGlyProAspAlaAspGly 1480
 DB 4413 TCCAGAGAGAGCGATCAGAGGAGAGGAGCAAGCCAGATGCTGCTCTGATGCTGACGGC 4472
 QY 1481 LysGluSerThrAlaIleGluLysValLeuLysAlaGluProGluIleLeuGluLeuGlu 1500
 DB 4473 AAGAGAGTACACAAATCGGAAAAGTCTCAAGCTGAACTGAGATCCTGGAACTTGAG 4532
 QY 1501 SerLysSerAsnLysIleValLeuAsnValIleGlnThrAlaValAspGlnPheAlaArg 1520
 DB 4533 AGTAAGAGCAACAAGATTGTGCTGAACGTCATTCAGACACGCCGTGACCAAGTTCGCAAGT 4592
 QY 1521 ThrGluThrAlaProGluThrHisAlaTyroAspSerGlnThrGlnValProAlaCysArg 1540
 DB 4593 ACAGAAACAGCCCCGAAACTCATGCTTATGATTCACAGACCCAGGTTCTCGCAATGCAG 4652
 QY 1541 LeuAspSerArgGluProAsnArgCysTrpThrLysMetLysAspAlaLysMetLysHis 1560
 DB 4653 GCTGACAGAGGAGGCCAACAGATGCTGGACAAAATGAAGTTGCCAAGATGAACAC 4712
 QY 1561 ProValProGlnProArgGluAspLeuGlnValLeuThrValLeuGluAlaTrpAlaGln 1580
 DB 4713 CCAGTGGCCGAGCCAGAGAGGACTTCCAAGTCTCGACCGTCTGGAGGATGGGCTCAG 4772
 QY 1581 ProArgLysCysLeuProArgLeuGlnLeuLysAlaProValSerLys 1596
 DB 4773 CCTCGAAATGCTTGGCGCGTTCGAGTTGAAAGCGCGGTGTCAAAG 4820

RESULT 3

AC115647 178954 bp DNA linear HTG 13-JUL-2002
 LOCUS Rattus norvegicus clone CH230-171G16, *** SEQUENCING IN PROGRESS
 DEFINITION ***, 53 unordered pieces.
 ACCESSION AC115647
 VERSION AC115647.3 GI:21737529
 KEYWORDS HTG; HTGS PHASE1.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 178954)
 Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C.,
 Alsbrooks, S.L., Amaratunge, H.C., Are, J.R., Ayele, M., Banks, T.,
 Barbarta, J., Benton, J., Binage, K., Blankenburg, K., Bonnin, D.,
 Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,
 Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,
 Carron, T.F., Carter, M., Cavazos, S.K., Chacko, J., Chavez, D.,
 Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,
 Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
 Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,
 Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, K.H.,
 Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
 Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
 Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
 Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
 Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
 Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,
 Hernandez, O., Hodgson, A., Hognes, M., Holloway, C., Hollins, B.,
 Homs, P., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,
 Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
 Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
 Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
 Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louseged, H.,
 Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
 Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,

```

Massey, E., Mawhiney, E., McLeod, M. P., Meador, M., Mei, G., Meeker, M.,
Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
Nguyen, N., Nickerson, E., Nwokkwo, S., Ogun, M., Okunolu, G.,
Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
Peters, L., Pickens, R., Primus, E., Pu, L., Quiles, M., Ren, Y.,
Rivers, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savary, G.,
Scheerer, S., Scott, G., Shen, H., Shooshkat, N., Sison, I.,
Sodergren, E., Sonalke, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Syatek, A., Taber, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Umani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
Williams, G., Williamson, A., Wiczek, R., Wooden, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G. and Gibbs, R.
Unpublished
2 (bases 1 to 178954)
Worley, K. C.
Direct Submission
Submitted (22-MAR-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 178954)
Worley, K. C.
Direct Submission
Submitted (13-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 12, 2002 this sequence version replaced gi:21239989.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu

----- Project Information
Center project name: GPHP
Center clone name: CH230-171G16
----- Summary Statistics
Sequencing vector: plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 142023 bases at least Q40
Consensus quality: 146696 bases at least Q30
Consensus quality: 150215 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 53 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
1459: contig of 1459 bp in length
1460
1559: gap of unknown length
1560
2931: contig of 1372 bp in length
2932
3031: gap of unknown length
3032
4420: contig of 1389 bp in length
4421
4520: gap of unknown length
4521
5800: contig of 1280 bp in length
5801
5900: gap of unknown length
5901
7559: contig of 1659 bp in length
7560
7659: gap of unknown length
7660
8887: contig of 1228 bp in length
8888
8987: gap of unknown length
8988
10297: contig of 1310 bp in length
10298
10397: gap of unknown length
10398
12090: contig of 1693 bp in length
12091
12190: gap of unknown length

12191
13523
13523
13623
15623
15623
15630
17523
17541
17541
17641
19820
19820
19920
21261
21362
23786
23786
23886
2553
2563
27443
27443
27543
30416
30516
31785
31785
31885
34971
35071
36234
36234
36334
39659
39659
39769
42867
42867
46088
46088
46189
49027
49027
49128
49128
51580
51580
55193
55193
55293
58205
58205
60939
60939
61040
61040
64185
64185
64285
67777
67777
67877
72020
72020
72120
75148
75148
75248
78406
78406
82239
82239
82339
85154
85154
85254
87767
87767
87868
87868
90644
90644
90744
94014
94014
98294
98294
98394
101683
101683
101783
104238
104238
104337
108706
108706
108807
108807
113208
113208
113209
113309
117239
117239
117339

13523: contig of 1332 bp in length
13622: gap of unknown length
15623: contig of 2007 bp in length
15623: gap of unknown length
15630: gap of unknown length
17523: contig of 1811 bp in length
17541: gap of unknown length
17541: gap of unknown length
17641: contig of 2179 bp in length
19819: gap of unknown length
19819: gap of unknown length
19919: gap of unknown length
21261: contig of 1342 bp in length
21361: gap of unknown length
23785: contig of 2424 bp in length
23885: gap of unknown length
25562: contig of 1677 bp in length
25662: gap of unknown length
27442: contig of 1780 bp in length
27542: gap of unknown length
30415: contig of 2873 bp in length
30515: gap of unknown length
31784: contig of 1269 bp in length
31884: gap of unknown length
31970: contig of 3086 bp in length
34970: gap of unknown length
35070: gap of unknown length
36233: contig of 1163 bp in length
36333: gap of unknown length
39668: contig of 3335 bp in length
39768: gap of unknown length
42866: contig of 3098 bp in length
42966: gap of unknown length
46088: contig of 3122 bp in length
46188: gap of unknown length
49027: contig of 2839 bp in length
49127: gap of unknown length
51579: contig of 2452 bp in length
51679: gap of unknown length
55192: contig of 3513 bp in length
55292: gap of unknown length
58204: contig of 2912 bp in length
58304: gap of unknown length
60939: contig of 2635 bp in length
61039: gap of unknown length
64184: contig of 3145 bp in length
64284: gap of unknown length
67776: contig of 3492 bp in length
67876: gap of unknown length
72019: contig of 4143 bp in length
72119: gap of unknown length
75147: contig of 3028 bp in length
75247: gap of unknown length
78405: contig of 3158 bp in length
78505: gap of unknown length
82238: contig of 3733 bp in length
82338: gap of unknown length
85153: contig of 2815 bp in length
85253: gap of unknown length
87767: contig of 2514 bp in length
87867: gap of unknown length
90643: contig of 2776 bp in length
90743: gap of unknown length
94013: contig of 3270 bp in length
94113: gap of unknown length
98293: contig of 4180 bp in length
98393: gap of unknown length
101682: contig of 3289 bp in length
101782: gap of unknown length
104237: contig of 2455 bp in length
104337: gap of unknown length
108705: contig of 4369 bp in length
108806: gap of unknown length
113208: contig of 4402 bp in length
113308: gap of unknown length
117238: contig of 3930 bp in length
117338: gap of unknown length
122577: contig of 5239 bp in length

```

* 122578 122677: gap of unknown length
 * 122678 126779: contig of 4102 bp in length
 * 126780 126879: gap of unknown length
 * 126880 131807: contig of 4928 bp in length
 * 131808 131907: gap of unknown length
 * 131908 138315: contig of 6408 bp in length
 * 138316 138415: gap of unknown length
 * 138416 143597: contig of 5182 bp in length
 * 143598 143697: gap of unknown length
 * 143698 152755: contig of 9058 bp in length
 * 152756 152855: gap of unknown length
 * 152856 157897: contig of 5042 bp in length
 * 157898 157997: gap of unknown length
 * 157998 167260: contig of 9263 bp in length
 * 167261 167360: gap of unknown length
 * 167361 178954: contig of 11594 bp in length.

FEATURES

Location/Qualifiers

source 1..178954

Alignment Scores:

Pred. No.: 4, 88e-257 Length: 178954
 Score: 7461.00 Matches: 1488
 Percent Similarity: 99.20% Conservatives: 6
 Best Local Similarity: 98.80% Mismatches: 11
 Query Match: 92.42% Indels: 3
 DB: 2 Gaps: 0

US-09-902-432-4 (1-1596) x AC115647 (1-178954)

QY 89 ValGlyGlnArgGluSerGluAspValArgGluLysAspArgValGluGluMetAlaAla 108
 Db 174384 GTTGGACGGCAGAGT CAGAAAGATGTGAGAGAAAGAACCGAGTTGAAGAAATGGCGGCC 174443
 QY 109 AsnSerThrAlaValGluAspIleThrLysAspGlyGlnGluGluThrSerGluIle 128
 Db 174444 AACTCCACAGCTGTTGAAGATATCACAAAGATGGGACAGAGAGACATCAGAAATATT 174503
 QY 129 GluGlnIleProAlaSerGluAsnAsnValGluGluMetValGlnProAlaGluSerGln 148
 Db 174504 GAACAGATCCCTGCTTCAGAAAGCAATGTGTGAAGAAATGGTACAGCCCTGCTGAGTCCAG 174563
 QY 149 AlaAsnAspValGlyPheLysLysValPheLysPheValGlyPheLysPheThrValLys 168
 Db 174564 GCTAATGATGTGGCTTCAGAAAGATATTTAAATTTGTTGTTTAAATTCACGGTGAAG 174623
 QY 169 LysAspLysAsnGluLysSerAspThrValGlnLeuLeuThrValLysLysAspGluGly 188
 Db 174624 AAGGATAAAATGAAGATCAGATACCTGTCACACTACTCTGTCACAGAGGATGAAGGC 174683
 QY 189 GluGlyAlaGluAlaSerValGlyAlaGlyAspHisGlnGluProSerValGluThrAla 208
 Db 174684 GAAGGGGCAGAAGCCTCTGTGGAGCTGGAGACCCACCAGGAGCCAGTGTGGAGACTGCC 174743
 QY 209 ValGlyGluSerAlaSerLysGluSerGluLeuLysGlnSerThrGluLysGlnGluGly 228
 Db 174744 GTCGGAGAGTCAGCATCCCAAGAAAGTGAGCTGAAGCAATCCACAGAGAAGCAAGAGGC 174803
 QY 229 ThrLeuLysGlnGluGlnSerSerThrGluIleProLeuGlnAlaGluSerAspGlnAla 248
 Db 174804 ACCCTGAAGCAGACAGACAGCAGCAAGAAATCCCCCTTCAAGCCGAAATCTGATCAAGCG 174863
 QY 249 AlaGluGluAlaLysAspGluGlyGluLysGlnGlnSerThrGluLysGlnGluGly 268
 Db 174864 GCTCAGGAAGAAGCAAGATGAAGAGAGAGAAACCAAGAGAAAGGCCCAAGTCC 174923
 QY 269 ProGluSerProSerSerProValAsnSerGluThrThrSerSerPheLysPhePhe 288
 Db 174924 CCAGAAATCCCCGAGACGCCAGTCAGCAGTGAGACAAACATCTTCTTCAAGAAATTTCTTC 174983
 QY 289 ThrHisGlyTrpAlaGlyTrpArgLysLysThrSerPheLysLysSerLysGluAspAsp 308
 Db 174984 ACTCACGGTTGGGCGGCTGGCGCAGAGACCCAGCTTCAAGAAATCAAAGAGAGATGAT 175043

QY 309 LeuGluThrAlaGluLysArgLysGluGlnGluAlaGluLysValAspGluGluGluLys 328
 Db 175044 CTGGAACCTGCCGAGAGAGAAAGAGGAGCAAGAGGCGAGAAAAGTAGACGAGGAGAAAAG 175103
 QY 329 GluLysThrGluProAlaSerGluGluGlnGluProAlaGluAspThrAspGlnAlaArg 348
 Db 175104 GAAAACAGACAGACCGCTCGGAGGAGCAGAGCCGGCAGAGACACACAGACCAGGCCAGG 175163
 QY 349 LeuSerAlaAspTyrGluLysValGluLeuProLeuGluAspGlnValGlyAspLeuGlu 368
 Db 175164 TTGTCAGCAGACTACGAGAAGGTGGAGCTGCCCTTGGAGACCAGGTTGCTGACCTGGAG 175223
 QY 369 AlaSerSerGluLysCysAlaProLeuAlaThrGluValPheAspGluLysMetGlu 388
 Db 175224 GCATCGTCAGAGGAGAGTGTGCTCTCTTTGGCAACGAAAGTGTTCATGAGAAAGATGAA 175283
 QY 389 AlaHisGlnGluValValAlaGluValHisValSerThrValGluLysThrGluGluGlu 408
 Db 175284 GCCCACCMAGNAGTTGTTGCAGAGGTCCAGTGAGCACCCCTGGAGAGACAGAGGAGGAG 175343
 QY 409 GlnGlyGlyGlyGlyGluAlaGluGlyGlyValValValGluGlyThrGlyGluSerLeu 428
 Db 175344 CAGGAGGAGGAGAGAGGCTGAAGGGGGCGTGTGTGAGAGGAGACAGAGAGATCCTTG 175403
 QY 429 ProProGluLysLeuAlaGluProGlnGluValProGlnGluAlaGluProAlaGluGlu 448
 Db 175404 CCCCCTGAGAAACTGGCTGAGCCCCCAGGAGGTCCCCCAGGAAGCTGAGCCCTGCTGAGGAG 175463
 QY 449 LeuMetLysSerArgGluMetCysValSerGlyGlyAspHisThrGlnLeuThrAspLeu 468
 Db 175464 CTGATGAAGAGCAGAGAGATGTGTCTCTGGAGGAGACACACTCAACTGACAGACCTA 175523
 QY 469 SerProGluGluLysThrLeuProLysHisProGluGlyIleValSerGluValGluMet 488
 Db 175524 AGTCCTGAGAGAGAGAGCTGCCCCAACACCCAGAGAGCAATTCAGTGAGGTGGAGATG 175583
 QY 489 LeuSerSerGlnGluArgLysValGlnGlySerProLeuLysLysLeuPheSerSer 508
 Db 175584 CTGTCCTCTCAGGAAAGAAATCAAGGTACAGGGAAGTCCCTTGAAGAAACTCTTTCAGTAGC 175643
 QY 509 SerGlyLeuLysLysLeuSerGlyLysLysGlnLysGlyLysArgGlyGlyGlyAsp 528
 Db 175644 TCAGGCTTAAAGAAAGCTGTCTGGGAAGAGACAGAGGGGAAACGAGGAGTGGGGAGAC 175703
 QY 529 GluGluProGlyGluTyrGlnHisIleHisThrGluSerProGluSerAlaAspGluGln 548
 Db 175704 GAAGACCTGGAGAAATCCCAACACATTCACACCGAATCCCCAGAGAGTCTGTATGAGCAG 175763
 QY 549 LysGlyGluSerSerAlaSerSerProGluGluProGluGluThrThrCysLeuGluLys 568
 Db 175764 AAGGGAGAGAGCTCTGCTGCTGCTCCCCGAGGAGCCCTGAGGAGACCACTGTCTGGAGAAA 175823
 QY 569 GlyProLeuGluAlaProGlnAspGlyGluAlaGluGluGlyThrThrSerAspGlyGlu 588
 Db 175824 GGCCCTCGAAGACACCCAGGATGGGGAAGCTGAGGAAGGAGCTGAGAGAGGCTCAAGAGCGCC 175883
 QY 589 LysLysArgGluGlyIleThrProTyrAlaSerPheLysLysMetValThrProLysLys 608
 Db 175884 AAGAAAGGAGAGGAGATCACTCCCTGGGCAATCCCTTCAAAAAGATGGTGACACCCCAAGAAA 175943
 QY 609 ArgValArgArgProSerGluSerAspLysGluGluLeuGluLysValLysSerAla 628
 Db 175944 CGGGTCCGAAGACCTTCTCAGAGTGACAAGGAGGAGAGCTGGAGAAGGTCAAGAGCGCC 176003
 QY 629 ThrLeuSerSerThrAspSerThrValSerGluMetGluAspGluValLysThrValGly 648
 Db 176004 ACCTTCTCTCCACTGATAGCAGTGTGAGAAATCCAGAGATGAAGTCAAAACTGTGTGT 176063
 QY 649 GluGluGlnLysProGluGluProLysArgValAspValAspThrSerValSerTrpGluAla 668
 Db 176064 GAGGAACAAAGCCAGAGGAGAACCAAGCGTAGGGTGGATACTTTCAGTGTCTTGGGAAGCA 176123
 QY 669 LeuIleCysValGlySerSerLysLysArgAlaArgLysAlaSerSerSerAspGlu 688

Db 176124 CTGATTTGTGTCGATCATCCAGAAAGAGGAGGAGGCGATCCTTTCAGATGTGAA 176183
OY |||||
OY 689 G1yG1yPrOArGThLeuG1yG1yAspSerH1sArG1aG1uG1u1aSerLysAspLys 708
Db 176184 GGAGGGCCCAAGGACACTGGGAGGGGACAGTCAACAGCAGAGAGAGGCCAGCAAGCAAA 176243
OY |||||
OY 709 G1u1aG1yThrAspAlaValProAlaSerThrG1nG1uG1nAspG1n1aG1nG1ySer 728
Db 176244 GAAGCGGAAACAGACGCTGTTCTGCGACAGCCAGAGAGCAGAGCCAGCCAGCAAGAAAT 176303
OY |||||
OY 729 SerSerProG1uProAlaG1ySerProSerG1uG1uG1uG1uValSerThrTrpG1uSer 748
Db 176304 TCCTCAACCGAGCCAGCGGAGGAGCCCTTCCGAAGGGAGAGGTCTCCACTTGGAGGTGA 176363
OY |||||
OY 749 PheLysArgLeuValThrProArgLysSerLysSerLysLeuG1uG1uLysAlaG1u 768
Db 176364 TTTAAAGATTAGTCACTCCAGAAAATTCAGACTCAAACTGGAGAGAAAGCCGA 176423
OY |||||
OY 769 AspSerSerValG1uG1nLeuSerThrG1u1eG1uProSerArgG1uG1uSerTrpVal 788
Db 176424 GACTCTAGGTAGAGAGAGTTGTCCACTGAGATCGAAACGAGTAGAGAAATCTTGGGTT 176483
OY |||||
OY 789 SerL1eLysPheL1eProG1yArgArgLysLysArgAlaAspG1yLysG1nG1uG1n 808
Db 176484 TCCATTAAAGAAATTCATCCCGAGCGGAGAAAGGGCAGATGGAGAGCAAGCA 176543
OY |||||
OY 809 AlaThrValG1uAspSerG1yProValG1u1eAsnG1uAspAspProAsnValProAla 828
Db 176544 GCCACTGTGGAAAGACTCAGGGCCAGTGGAGATTAATGAGAGCACTCTATGTCAGGCC 176603
OY |||||
OY 829 ValValProLeuSerG1uTyraAsnAlaValG1uArgG1uLysMetG1u1aG1nG1yAsn 848
Db 176604 GTCGTGCTCTGTCTGAGATATGATGAGTGGAGAGGAGAGATGGAAACCCAGGGGAAT 176663
OY |||||
OY 849 ThrG1uLeuProG1nLeuLeuG1yAlaValTyraSerG1uG1uLeuSerLysThreLeu 868
Db 176664 GCGGAGCTGCCCCAGCTGTGGGGGTGTGTACGTGTCCGAGAGCTCGATAAGACTCTG 176723
OY |||||
OY 869 ValH1sThrValSerValAlaVal1leAspG1yThrArgAlaValThrSerValG1uG1u 888
Db 176724 GTCCACACTGTAGTGTGTCACATCATGTGGAGCCAGGGCAGTCACTGTCGAAAG 176783
OY |||||
OY 889 ArgSerProSerTrpL1eSerAlaSerValThrG1uProLeuG1uH1sThraG1yG1u 908
Db 176784 CGGTCTCTCTGATATCCCTTCCGTAAACAACCTCTGAACACACAGCGGAGAA 176843
OY |||||
OY 909 AlaMetProProValG1uG1uValThrG1uLysAspL1e1eAlaG1uG1uThrProVal 928
Db 176844 GCCATGCCACTGTGTGAAGAGTCACTGAAGAAAGATCATTCAGAAAGAACTCTGTG 176903
OY |||||
OY 929 LeuThrG1nThreLeuProG1uG1yLysAspAlaH1sAspAspMetValThrSerG1uVal 948
Db 176904 CTCACCCAGACCTTACAGAGGTAAAGATGCCCATGACAGCATGTGTACACAGTGAAGT 176963
OY |||||
OY 949 AspPheThrSerG1uAlaValThra1aThrG1uThrSerG1uAlaLeuArgThrG1uG1u 968
Db 176964 GATTTCACCTCAGAAAGCTGTGACAGCCACAGAACTCAGAGGCTCTCCGTACTGAAGA 177023
OY |||||
OY 969 ValThrG1uAlaSerG1yAlaG1uG1uThrThraPheMetValSerAlaValSerG1nLeu 988
Db 177024 GTTACCGAAGCATCGGGGCCAGAAAGACACAGACATGTGTGCGCAATTTCCCGCTG 177083
OY |||||
OY 989 ThrAspSerProAspThrThrG1uG1uAlaThrProValG1nG1uValG1uSerG1yVal 1008
Db 177084 ATTGACTCCCGACACACCAAGAGAGCCACCCAGTTCAAGAGGTAAAGGGTGTGTG 177143
OY |||||
OY 1009 LeuAspThrG1uG1uG1uG1uArgG1nThrG1nAla1leuG1nAlaVal1aAspLys 1028
Db 177144 CTAGATACAGAAAGAGAGGCGCCAGACGACGAGCCATCTCCAAAGCCGTGAGACAAAG 177203
OY |||||
OY 1029 ValLysG1uG1uSerG1nValProAlaThrG1nThraValG1nArgThrG1ySerLysAla 1048
Db |||||

Db 177204 GTGAAGAGAGTCCAGGTCTCTGCAACCAGACTGTGCAAGAAACGGGCTCAAAAGCA 177263
OY |||||
OY 1049 LeuG1uLysValG1uG1uValG1uG1uAspSerG1uValLeuAlaSerG1uLysG1uLys 1068
Db 177264 CTGGAGAGGTTGAGAGAGGTAGAGAGAGACTCCGAAGTGTGCTTCCGAGAAAGAGAG 177323
OY |||||
OY 1069 AspValMetProLysG1yProValG1nG1uAlaG1yLysAlaG1uH1sLeuAlaG1nG1ySer 1088
Db 177324 GACGTTATGCGGAAAGGACCCGTCCAGAAAGCTGAGCTGAGTCATTTGCAAGGGCTCT 177383
OY |||||
OY 1089 GluThrG1yG1nAlaThrProG1uSerLeuG1uValProG1uValThra1aAspValAsp 1108
Db 177384 GAGACTGACAGCTTACTCCAGAGAGCTTGAAGTTCTTGAAGTCAACAGCATGTAGAC 177443
OY |||||
OY 1109 H1sValAlaThrCysG1nVal1leLysLeuG1nG1uLeuMetG1uG1nAlaValAlaPro 1128
Db 177444 CATGTGCGCACGTGCGAGATTATCAAGTCTCAGAGCTGATGAAACAGCGCGTGGCCCT 177503
OY |||||
OY 1129 GluSerSerG1uThrLeuThraAspSerG1uThraArgLysSerThrProLeuAlaAspSer 1148
Db 177504 GAGTCATCCGAACCTTGAACAGACAGTGAACAATGGACACACTCCCTTGAAGATTCA 177563
OY |||||
OY 1149 AspThra1aAspG1yThrG1nG1uAspG1uThrL1eAspSerG1nAspSerLysAlaThr 1168
Db 177564 GACACTGCGATGGACACACAGCAAGATGAACATTGACAGCCAGAGCATGTAAAGCAGCT 177623
OY |||||
OY 1169 AlaAlaValArgG1nSerG1nValThrG1uG1uAla1aThra1aG1nLysG1uG1u 1188
Db 177624 GCAGCTGCAGGACGTCAACAGTCAACAGAAAGCGCGCTATCGTCAAGAAAGAGAG 177683
OY |||||
OY 1189 ProSerThrLeuProAsnAsnValProAlaG1nG1uG1uH1sG1yG1uG1uProG1yArg 1208
Db 177684 CCTTCGACACTCACTTAATTAATGTTCCAGCCAGAGAAACATGGGAAACCAAGGAAGA 177743
OY |||||
OY 1209 AspValLeuG1uProThrG1nG1uG1uLeuThra1aAlaValProValLeuAlaLys 1228
Db 177744 GATGTTCTTGAACCTTACACAGCAAGAGCTTGTGCTGACGCGTGTGCCCTTCTGGCAAG 177803
OY |||||
OY 1229 ThrG1uValG1yG1nG1uG1yG1uValAspTrpLeuAspG1yG1uLysValLysG1uG1u 1248
Db 177804 ACTGAGTGGGTCAAGAGGGTGAAGTGTGACTGTTGATGAGAAAGTCAAGAAAGCA 177863
OY |||||
OY 1249 G1nG1uValPheValH1sSerG1yProAsnSerG1nLysAla1aAspValThrTyraP 1268
Db 177864 CAGAGGCTGTTGTACACTCTGGAACCCAAAGTCAAAAGGCTGCTGATGATCAATATAC 177923
OY |||||
OY 1269 SerG1uValMetG1yValAlaG1yCysG1nG1uLysG1uSerThrG1uValG1nSerLeu 1288
Db 177924 AGTGAAGTGAATGGAGTGGCCGGGTGTCAAGAAAGAGAGTACTGAAGTGCAGAGTCTT 177983
OY |||||
OY 1289 SerLeuG1uG1uG1yG1uMetG1uThraAspValG1uLysG1uLysArgG1uThrLysPro 1308
Db 177984 AGCTCGAGAGAGGAGAGATGAGAACTGACGTTGAAAAGGAAAGGAGCAAGCAAGCA 178043
OY |||||
OY 1309 G1uG1nValSerG1uG1yG1uG1nG1uThra1aAlaProG1uH1sG1yG1yThrTy 1328
Db 178044 GAGCAAGTGAAGTGAAGAGGTGAGAGAGAAACAGCGCTCTCTGAGCAAGAAAGCACTTAC 178103
OY |||||
OY 1329 G1yLysProValLeuThreLeuAspMetProSerSerG1uArgG1yLysAlaLeuG1ySer 1348
Db 178104 GGGAAAGCAAGTCTGACACTTGACATGCCCCAGCTCAGAGAGGGGAGAGCACTGGGAAAGC 178163
OY |||||
OY 1349 LeuG1yG1ySerProSerLeuProAspG1nAspLysAlaG1yCys1leG1uValG1nVal 1368
Db 178164 CTTGAGAGGAGCCCTTCTCTCCAGACCAAGCAAGAGCGGTGTGATGAGGTTCAAGTT 178223
OY |||||
OY 1369 G1nSerLeuAspThrThra1aThrG1nThra1aG1uAlaValG1uLysVal1leG1uThr 1388
Db 178224 CAAGCCTGAGACAAACAGTCACTCAAAACAGCAGAAAGCTGTGGAAAAGTCAATAAAGCG 178283
OY |||||
OY 1389 ValVal1leSerG1uThrG1yG1uSerProG1uCysValG1yAlaH1sLeuLeuProAla 1408
Db 178284 GTTGATTTTCAGACACAGGTGAAGTCCAGAGTGTGATGATGACACTTATTACAGCT 178343
OY |||||

```
QY 1409 GluLysSerSerAlaThrGlyHisTrpThrLeuGlnHisAlaGluAspThrValPro 1428
|||
Db 178344 GAGAAGTCTCTCAACGGGTGGCCACTGGACTCTTCAGCATGCAGAGACACGGTACCC 178403
|||
QY 1429 LeuGlyProGluSerGlnAlaGluSerIleProIleIleValThrProAlaProGluSer 1448
|||
Db 178404 CTGGGGCTCAGTCTCAGGCAGAAATCCATCCCAATCATAGTAATCTCTCTCTGAAGC 178463
|||
QY 1449 ThrLeuHisProAspLeuGlnGlyGluIleSerAlaSerGlnArgGluArgSerGluGlu 1468
|||
Db 178464 ACCCTACATCTGACCTACAGGAGAAATAGCGCATCCAGAGAGCGCATCAGAGAA 178523
|||
QY 1469 GluAspLysProAspAlaGlyProAspAlaAspGlyLysGluSerThrAlaIleGluLys 1488
|||
Db 178524 GAGGACAAGCCAGATGCTGCTCTGATGCTGACGGCAAGGAGAGTACAGCAATCGAAAA 178583
|||
QY 1489 ValLeuIleAlaGluProGluIleLeuGluLeuGluSerLysSerIleValLeu 1508
|||
Db 178584 GTCTCAAGGCTGAACCTGAGATCTCTGAACTTGAGAGTAAAGACAAGATTGTGCTG 178643
|||
QY 1509 AsnValIleGlnThrAlaValAspGlnPheAlaArgThrGluThrAlaProGluThrHis 1528
|||
Db 178644 AACGTATTCAGACAGCGGTGACAGTTCGCAGTACAGAAACAGCCCCGAACTCAT 178703
|||
QY 1529 AlaTrpAspSerGlnThrGlnValProAla-CysArgLeuAspSerArgGluProAsnAr 1548
|||
Db 178704 GCTTATGATTCACAGACCCAGGTTCTGCAATGCAGCTCAGACGACGAGGAGCCCAACAG 178763
|||
QY 1548 CysTrpThrLysMetLysAspAlaLysMetLysHisProValProGlnProArgGluAs 1568
|||
Db 178764 ATGCTGGACAAAATGAAGTTCACAGATGAAACACCCAGTCCCGCAGCCACAGAGAG 178823
|||
QY 1568 pLeuGlnValLeuThrValLeuGluAlaTrpAlaGlnProArgLysCysLeuProArgLe 1588
|||
Db 178824 CTTGCAAGTCTCAGCCTTCTTGAGCA-TGGGCTCAGCCTCC-GAATGCTTGGCCGGCT 178881
|||
QY 1588 uGlnLeuAlaPro 1593
|||
Db 178882 GCANNTGAAAGCGCG 178897
|||
RESULT 4
AB020886 6195 bp mRNA linear ROD 14-APR-1999
LOCUS Mus musculus mRNA for SSeCKs, complete cds.
DEFINITION AB020886
ACCESSION AB020886
VERSION AB020886.1 GI:4589831
KEYWORDS SSeCKs.
SOURCE Mus musculus (strain: BALB/c, isolate: 200) 9-12 weeks male testis
tissue lib: Clontech mouse testis library cDNA to mRNA, clone: 133.5.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (sites)
Okita, K., Kitamura, H., Iwanaga, T., Morimatsu, M. and Saito, M.
Isolation and characterization of mouse SSeCKs
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 6195)
Okita, K., Kitamura, H., Saito, M. and Morimatsu, M.
Direct Submission
JOURNAL Submitted (08-DEC-1998) Keisuke Okita, School of Veterinary
Medicine, Hokkaido University, Laboratory of Biochemistry; N18, W9,
Kita-ku, Sapporo, Hokkaido 060-0818, Japan
(E-mail: v080008@vet.ec.hokudai.ac.jp, Tel: 81-11-706-5206,
Fax: 81-11-757-0703)
COMMENT Sequence updated (10-Mar-1999).
FEATURES
source
1. .6195
/organism="Mus musculus"
/strain="BALB/c"
/isolate="200"
/db_xref="taxon:10090"
/clone="133.5"
```

Db 320 GTCAATGGTGTAGCTGAACAGAAAGATGTCCAGTCCAAAGAGAAAAGCCAGATGGGCAA 379
 QY 81 GtuglgluValaIAspGluAspValGlyGlnArgGluSerGluAspValArgGlnuys 100
 Db 380 GAGGAAAGATGACTGCTTGAAGATGTTGACAGAGAGATCAGAAAGATGTGAAGAAAAA 439
 QY 101 AsparGValGluGluMetAlaIAsnSerThrAlaValGluAspLLeuThrLysAspGly 120
 Db 440 GACCAGCTAAAGAAATGCGACCCAGTTCCACAGTTGTTGAAGATATTCAAAGGCGAG 499
 QY 121 GlnGluGluThrSerGluLeuLeuGluGluGlnLeuProAlaSerGluAsnAsnValGluGlu 140
 Db 500 CAGGAGGAAACCCGAAATTAATCGAAGAGATCCCTGCTTCAGAGAGCAATGGAGAA 559
 QY 141 MetValGlnProAlaGluSerGlnAlaAsnAspValGlyPheLysLysValPheLysPhe 160
 Db 560 ATGGCGCAGGCGCTGAGTCCCAAGCTAATGACGTGCGCTTCAAGAAAGTATTAAATT 619
 QY 161 ValGlyPheLysPheThrValLysLysAspLysAsnGluLysSerAspThrValGlnLeu 180
 Db 620 GTTGGTTTAAATTCACGCGTGAAGAGATTAATAAAGCAAAAGTCAGATACCGTCCAGCTA 679
 QY 181 LeuThrValLysLysAspGluGluGlyValGluAlaSerValGlyAlaGlyAspHis 200
 Db 680 CTCACGTGTCAGAGAGATGAAGCGGAGGCGCAGAGCCCTCGGTGAGCAGAGACAC 739
 QY 201 GlnGluProSerValGluThrAlaValGlyGluSerAlaSerLysGluSerGluLeuLys 220
 Db 740 CAGAGCCCGGAGGTGAGAC---GTCGGCGAATCAGCATCCAAAGAAAGTACGTGAAG 796
 QY 221 GlnSerThrGluLysGlnGluGlyThrLeuLysGlnGluGlnSerThrGluLeuPro 240
 Db 797 CATCCACAGAGAGAGCAAGAGAGCCCTGAAGCAGCAGAGCAGCAGCAAGAAATTCCC 856
 QY 241 LeuGlnAlaGluSerAspGlnAlaAlaGluGluGluAlaLysAspGluGluGluGluLys 260
 Db 857 CTTCAAGCCGAACTGTCAGAGGACCGAGGAGAGAGCCCAAGATGAGAGAAAGAAC 916
 QY 261 GlnGluLysGluProThrLysSerProGluSerProGluSerProValAsnSerGluThr 280
 Db 917 CGAGAGAAAGAACTTCCAAAGCCCTTGAATCTCCGACAGCCCTGTCAGCAATGAGACA 976
 QY 281 ThrSerSerPheLysLysPhePheThrHisGlyTyrAlaGlyTyrPheGlyLysThrSer 300
 Db 977 ACATCTTCTTCAAGAAATCTTCACTCAAGCGCTGGCGCTGGCGCAAGAGACAG 1036
 QY 301 PheLysLysSerLysGluAspLeuGluThrAlaGluLysArgLysGluGluGluAla 320
 Db 1037 TTCAGAAACCAAGAGAGATGATCTGAAACTTCCAGAGAGAGAAAGAGAGAGGCT 1096
 QY 321 GlnLysValAspGluGluGluLysGluLysThrGluProAla---SerGluGluGlnGlu 339
 Db 1097 GAAAGAAAGTGAACAG 1156
 QY 340 ProAlaGluAspThrAspGlnAlaArgLeuSerAlaAspTyrGluLysValGluLeuPro 359
 Db 1157 CCGGAGAGAGAGCACAAGCAGGCGCAGGTTGTCAAGCCAGTATGAGAGAGGAGTGGCT 1216
 QY 360 LeuGluAspGlnValGlyAspLeuGluAlaSerSerGluGluLysCysAlaProLeuAla 379
 Db 1217 TTGGAAGACAGGCGTGTGACTGAGAGGCACTGTGC---GAGAGAGTGTGCTCTTGGCA 1273
 QY 380 ThrGluValPheAspGluLysMetGluAlaHisGlnGluValValAlaGluValHisVal 399
 Db 1274 ACGGAAGTGTATATAG 1333
 QY 400 SerThrValGluLysThrGluGluGlnGluGlyGlyGlyGluAlaGluGluGlyVal 419
 Db 1334 AGCAGCGTGAAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1390
 QY 420 ValValGluGluGlyThrGluGluSerLeuProProGluLysLeuAlaGluProGlnGluVal 439
 Db 1391 GTGGTGAAGAGATCGGAGAGATCTTGGCCCTCGAGAAAGTGGCTGAGACCCAGAGAGT 1450

QY 440 ProGlnGluAlaGluProAlaGluGluLeuMetLysSerAlaArgGluMetCysValSerGly 459
 Db 1451 CCCCAG 1510
 QY 460 GlyAspHisThrGluGlnLeuThrAspLeuSerProGluGluLysThrLeuProLysHis 479
 Db 1511 GGTGACCACTACTCAGCTGACAGATCTTAAGTCTTAAGAGAGAGAGAGAGAGAGAGAG 1570
 QY 480 GlnGlyLeuValSerGluValAlaGluMetLeuSerSerGlnGluArgGlyLeuValGlnGly 499
 Db 1571 GAAGGCAATTGTCAGTGAAGTGAAGTGTGCTGCTCCTCAGAGAGAGATCAAGTACAGGGA 1630
 QY 500 SerProLeuLysLysLeuPhePheSerSerSerGlyLeuLysLysLeuSerGlyLysGln 519
 Db 1631 AGTCCCTGAAGAGAGCTTCTTCAAGCAGGTTCCGAGCTTAAAGAGAGCTTCCGAGAGAG 1690
 QY 520 LysGlyLysArg---GlyGlyGlyGlyAspGluGluProGluGlyGlyTyrGlnHis 538
 Db 1691 AAGGAG 1750
 QY 539 ThrGluSerProGluSerAlaAspGluGlnLysGlyGluSerSerAlaSerSerProGlu 558
 Db 1751 ACCGAGTCCCAAGAGAGTCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1810
 QY 559 GlnProGluGluThrThrCysLeuGluLysGlyProLeuGluAlaProGlnAspGlyGlu 578
 Db 1811 GAGCCCGAGAGAGATCGCTGTCTGGAGAGAGGCGCATCGAGAGAGAGAGAGAGAGAG 1870
 QY 579 AlaGluGluGluGlyThrThrSerAspGlyGluLysLysArgGlyGlyLeuThrProThrAla 598
 Db 1871 GGTGAG 1930
 QY 599 SerPheLysLysMetValThrProLysLysArgValAlaGlyArgProSerGluSerAspLys 618
 Db 1931 TCCTTCAAAAAGATGTGATACACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1990
 QY 619 GlnGluGluLeuGluLysValLysSerAlaThrLeuSerSerThrAspSerThrValSer 638
 Db 1991 GAAAGAGAGCTGATTAAGATCAAGAGTCCACCTTGTCTCCACAGAGAGAGAGAGAG 2050
 QY 639 GlnMetGluAspGluValLysThrValGlyGluGluGlnLysProGluGluProLysArg 658
 Db 2051 GGAATGACAGAGATGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2110
 QY 659 ArgValAspThrSerValSerTyrGluAlaLeuLysCysValGlySerSerLysArg 678
 Db 2111 AGGTGTGATCTTCACTGCTCTTGGAGGCGCTTGTGTGATGCTGCTCAAGAGAGAG 2170
 QY 679 AlaArgLysAlaSerSerSerSerAspGluGlyGlyProArgThrLeuGlyGlyAspSer 698
 Db 2171 GCGAGAGAGAGCATCTCTTCAAGATGATGAAGAGAGAGAGAGAGAGAGAGAGAGAG 2230
 QY 699 HisArgAlaGluGluAlaSerLysAspLysGluAlaGlyThrAspAlaValProAlaSer 718
 Db 2231 CACAGAGCGGAG 2284
 QY 719 ThrGlnGluGluAspGluAlaGlnLysSerSerSerProGluProAlaGlySerProSer 738
 Db 2285 ACCCAG 2344
 QY 739 GlnGlyGluGlyValSerThrTyrPheLysPheLysArgLeuValThrProArgLysLys 758
 Db 2345 GAAGGAG 2404
 QY 759 SerLysSerLysLeuGluGluGluLysAlaGluAspSerSerValGlnGlnLeuSerThrGlu 778
 Db 2405 TCCAAGTCAAAAAG 2464
 QY 779 IleGluProSerArgGluGluSerTyrValSerIleLysLysPheIleProGlyLysArg 798
 Db 2465 ATCGAAGCAAGTAG 2524

QY 799 LyeIysArgAlaAspGlyLysGlnGluAlaThrValGluAspSerGlyProValGlu 818
DB 2525 AAAAAAGGCGAGATGGAGACGAAGAAGACAGCGCCCGCTTGAAGACTCGGGGCCAGGAG 2584
QY 819 IleAsnGluAspAspProAsnValProAlaValValProLeuSerGluTyrAsnAlaVal 838
DB 2585 ATCAATGAGGACGACCCCGAGCTCCAGCTGTGTGGCTCTGTCTGATGATGCGGTA 2644
QY 839 GluArgGluLysMetGlu-----AlaGlnGlyAsnThrGluLeuProGlnLeuLeu 855
DB 2645 GAGAGAGAAGAGCTGGAAAGCGCAGCGAGCTCAGGAGAACGTGGAGCTGCCCGAGTGAAG 2704
QY 856 GlyAlaValTyrValSerGluGluLeuSerLysThrLeuValHisThrValSerValAla 875
DB 2705 GGGGCTGTGTATGTCTGAGGAGCTTAGTAAGACTCTGGTTACACCGTGAAGTGTCCGG 2764
QY 876 ValIleAspGlyThrArgAlaValThrSerValGluGluArgSerProSerTrpIleSer 895
DB 2765 GTCAATTGATGGACACGAGCGAGTCAACAGTCCCGAAGAGCGGTCCCTTCTGTGGATATCT 2824
QY 896 AlaSerValThrGluProLeuGluHisThrAlaGlyGluAlaMetProProValGluGlu 915
DB 2825 GCTTCCATGACAGAACTCTTGTAGCACGACGAGGGAGTGGCCACACCGCTGTGTGGAGAG 2884
QY 916 ValThrGluLysAspIleIleAlaGluGluThrProValLeuThrGlnThrLeuProGlu 935
DB 2885 GTCACTGAAAAGACATCATCTGCAAGAACACTCTCCATCGCCCACTTTACCAAGGG 2944
QY 936 GlyLysAspAlaHisAspAspMetValThrSerGluValAlaAspPheThrSerGluAlaVal 955
DB 2945 GGCNAGATGCCCATGACGATAGTCAACAGTCCAGTGAAGTGAATTTACTCTAGAAGCAGTG 3004
QY 956 ThrAlaThrGluThrSerGluAlaLeuArgThrGluGluValThrGluAlaSerGlyAla 975
DB 3005 ACAGCCGACAGAAACACAGAGGGCGCTCCGCGCTGAAGAACTTACCGAAGCATCAGGGGCA 3064
QY 976 GluGluThrThrAspMetValSerAlaValSerGlnLeuThrAspSerProAspThrThr 995
DB 3065 GAAGAGACCCACAGATGTGTCTGCAAGTTTCCAGAGTGTCCGACTCCCGCGACACACA 3124
QY 996 GluGluAlaThrProValGlnGluValGluSerGlyValLeuAspThrGluGluGlu 1015
DB 3125 GAGNAGCCACCCAGTTACGAGGTAGAGGTGGCATGTAGATACGGAAGAACAGGAG 3184
QY 1016 ArgGlnThrGlnAlaIleLeuGlnAlaValAlaAspLysValLysGluGluSerGlnVal 1035
DB 3185 CGCCAGACGAGCGCGTCTCCAGCGCTTGCACAGAAAGTGAAGAGGACTCCCAGTG 3244
QY 1036 ProAlaThrGlnThrValGlnArgThrGlySerLysAlaLeuGluLysValGluGluVal 1055
DB 3245 CCTGCCAACCCAGACTCTGCAGAGACGAGCGCCGAAAGCACTGGAGAAGTGGAGAGGTA 3304
QY 1056 GluGluAspSerGluValLeuAlaSerGluLysAspValMetProLysGlyPro 1075
DB 3305 GAGAGAGACTCCGAGGTGTGGCTTACCGAAGAGAAGAGATGTGTGCGGAAGGACCC 3364
QY 1076 ValGlnGluAlaGlyAlaGluHisLeuAlaGlnGlySerGluThrGlyGlnAlaThrPro 1095
DB 3365 GTGCAGGAAGCTGAAACTGAGCATCTTGCACAGGGCTCCGAGACTGTACAGGCTACCCCA 3424
QY 1096 GluSerLeuGluValProGluValThrAlaAspValAspHisValAlaThrCysGlnVal 1115
DB 3425 GAGAGCCTTGAAGTTCTGAAGTCACAGAGGATGTAGACGCTGACACCATGCGCAGGTT 3484
QY 1116 IleLysLeuGlnGlnLeuMetGluGlnAlaValAlaProGluSerSerGluThrLeuThr 1135
DB 3485 ATCAAGCACCAGCAGCTGATGGAAACAGCGTGTGGCCCTCCTGAGTCATCTGAACCTTGACA 3544
QY 1136 AspSerGluThrAsnGlySerThrProLeuAlaAspSerAspThrAlaAspGlyThrGln 1155
DB 3545 GACAGTGAGACNAATGAAGTACTCCCTCGCAGATTACAGACTCCCAACGGGACACAG 3604
QY 1156 GlnAspGluThrIleAspSerGlnAspSerLysAlaThrAlaValArgGlnSerGln 1175

DB 3605 CAAGACGAGACCGTTTCACAGCCAGGACAGTAATGCCATTGCGCCCGCTCAAGCAGTCACAG 3664
QY 1176 ValThrGluGluGluAlaAlaThrAlaGlnLysGluGluProSerThrLeuProAsnAsn 1195
DB 3665 GTCACTGAAGAGGAGGAGCGAGCTGCTGCAGCGAGGGGCTTCAACA---CCATCTAGT 3721
QY 1196 ValProAlaGlnGluGluHisGlyGluGluProGlyArgAspValLeuGluProThrGln 1215
DB 3722 TTTCCAGCCCGAGAAACACACAGGAAAAACAGGAAGGATGTTCTAGAACCCACA--- 3778
QY 1216 GlnGluLeuThrAlaAlaValProValLeuAlaLysThrGluValGlnGluGly 1235
DB 3779 CAAGCCCTGCTGCTCGGGGAGTGCCTATTCTGGCAAGGCTGAGGTGGGTCAAGAGGT 3838
QY 1236 GluValAspTrpLeuAspGlyGluLysValLysGlu-----GluGlnGlu 1250
DB 3839 GAGCTGCCCGACTTTCATGGAGAAAAGTCAAGACGGACAGTGTGTTAAAGAACTGGAG 3898
QY 1251 ValPheValHisSerGlyProAsnSerGlnLysAlaAlaAspValThrTyrAspSerGlu 1270
DB 3899 GTGCCTGTGCACACTGGACCCCAACAGTCAAAAGACTGCTGACTTGACACGTCAGAGTAA 3958
QY 1271 ValMetGlyValAlaGlyCysGlnGluLysGluSerThrGluValGlnSerLeuSerLeu 1290
DB 3959 GTAATGGAAGTGGCCAGATGTCAAGAAACCTGAGAGTAATGAAGAACAGAGTATTAGCCG 4018
QY 1291 GluGluGlyGluMetGluThrAspValGluLysGluLysArgGluThrLysProGluGln 1310
DB 4019 GAGAAAGAGAGATGGGAACCGACGTTGAAAGAGGAGAAACAGAGACCAAGACAGAGCAA 4078
QY 1311 ValSerGluGluGlyGluGlnGluThrAlaAlaProGluHisGluGlyThrTyrGlyLys 1330
DB 4079 GCCAGTGAAGAACATCAGCAGGAAACAGCTGCTCTCTGAGCATGAAGGAACCCACCTAAG 4138
QY 1331 ProValLeuThrLeuAspMetProSerSerGluArgGlyLysAlaLeuGlySerLeuGly 1350
DB 4139 CCAGTCTTCACAGCTACATGCCTCACTCAGAGAGGGAAAGGCATCGGCGACCTTGAA 4198
QY 1351 GlySerProSerLeuProAspGlnAspLysAlaGlyCysIleGluValGlnValGlnSer 1370
DB 4199 GGAAGCCCTTCTCTCCACAGCCACGAAGCAAGCAGATTGCATAGAGGTTCAAGTTCAAAGC 4258
QY 1371 LeuAspThrThrValThrGlnThrAlaGluAlaValGluLysValIleGluThrValVal 1390
DB 4259 TCAGACACACCACTCACTCAAAACACCGAGCTGTGAAAAAGTCCGAAGAACTGTGGCA 4318
QY 1391 IleSerGluThrGlyGluSerProGluCysValGlyAlaHisLeuLeuProAlaGluLys 1410
DB 4319 ACTTCAGAGATGGATGAAAGTTTGGAGTGTGCAGGTGCGCAATCATATTACAGCTGAGAAG 4378
QY 1411 SerSerAlaThrGlyGlyHisTrpThrLeuGlnHisAlaGluAspThrValProLeuGly 1430
DB 4379 CTCTCCGAAACCGGTGGCTACGGGACTTTCAGCATGGAGAGGACACCGTGCCTCCAGGG 4438
QY 1431 ProGluSerGlnAlaGluSerIleProIleIleValThrProAlaProGluSerThrLeu 1450
DB 4439 CCTGAGTCTCAGSCAGAGTCCATCCCAATAATAGTAACCTCTCTCTCTGAAAGCATCCTTA 4498
QY 1451 HisProAspLeuGlnGlyGluIleSerAlaSerGlnArgGluArgSerGluGluGluAsp 1470
DB 4499 CATTCGACCTTCAAGAGAAAGTGAAGCATCCCAAGAAACAGAGATCAGATGAAGATAAC 4558
QY 1471 LysProAspAlaGlyProAspAlaAspGlyLysGluSerThrAlaIleGluLysValLeu 1490
DB 4559 AAGCCAGATGCTGGTCTCTGATGCTGCCGCAAGGAGAGTGCAGAGAGAGAAAAATCCTC 4618
QY 1491 LysAlaGluProGluIleLeuGluLeuGluSerLysSerAsnLysIleValLeuAsnVal 1510
DB 4619 AGGGCTGAACCTGAGATCTTGGNACTTGAGATGAAGCAATPAAGATTGTCCAGAGTGTCTC 4678
QY 1511 IleGlnThrAlaValAspGlnPheAlaArgThrGluThrAlaProGluThrHisAlaTyr 1530

Db 4679 ATCCGACAGCCGTCGACCGACGTTTGACCTACAGAAACAGCCCGGAAACCCAGCGCTTCT 4738
QY 1531 AspSerGlnThrGlnValProAla-CysArgLeuAspSerArgGluProAsnArgCysTr 1550
Db 4739 GATTACAGCAATCAGATCTCTGTGATGACGCTGACGACGAGGACACACAGATGCTG 4798
QY 1550 PThrIysMetLys-----AspAlaIysMetLysHisLeuProValProGlnPr 1565
Db 4799 GACAAAGATGAAAGACGACCTTCAAGTCTCCCGCAAGATGAAACACTGACTGCGGAGCC 4858
QY 1565 cArgGluAspLeuGlnValLeuThrValLeuGlnAlaIatPAlaGlnProArgLysCysLe 1585
Db 4859 CAGGAAGACCTTGCGGTTCTGATAGTTCTTGAAGCATACAGCAAGCTTCAGAAATGATC 4918
QY 1585 uProArgLeuGlnLeuLysAlaIleProValSerLys 1596
Db 4919 ACCACGCTTGACGTTGAAAGTGCCAGCTGCAAG 4952
RESULT 5
AF326228 6197 bp mRNA linear ROD 16-JUL-2001
LOCUS AF326228
DEFINITION Mus musculus PKC binding protein SSeCKs mRNA, complete cds.
ACCESSION AF326228
VERSION AF326228.1 GI:13236403
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 6197)
Camus,A., Mesbah,K., Rallu,M., Babinet,C. and Barra,J.
Gene trap insertion reveals two open reading frames in the mouse
SSeCKs gene: the form predominantly detected in the nervous system
is suppressed by the insertion while the other, specific of the
testis, remains expressed
Mech. Dev. 105 (1-2), 79-91 (2001)
21322683
11429284
2 (bases 1 to 6197)
Camus,A., Mesbah,K., Babinet,C. and Barra,J.
Direct Submission
Submitted (06-DEC-2000) Immunologie, Institut Pasteur, 25 rue du
Docteur Roux, Paris 75015, France
Location/Qualifiers
1..6197
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="10"
140..5194
/codon_start=1
/product="PKC binding protein SSeCKs"
/protein_id="AAK16150.1"
/translation="MGAGSSTRORSPEOPASDTPSELLESGHPAASGAAGPAD
ADPATLPEOKQLSANGVAAQEBVHVOEEOQOEVEYTVEDQGESEBVEKDR
AKEMASSTVVEDITKDEEETPEITIEQIPASEVVEEAQVABEQANDVKKAFK
VGFKTVKDKNEKSDTVQLTVKKBDEGASVAGQADGENNEKEPTPLPESVSG
KOSTKEOEGTLKQASSTEIPLQASGGQTEBAADKGENNEKEPTPLPESVSG
NETSSPFKFTFHGAWGMRKSTSPFKPKEDDLETSEKKEQAEKAVDEEGEETPAP
AEOPAGCTDQARLSADYEKELPLEDQEGDLSEKCAFLATVEDEKTEADQEV
VAEVAVSTERTKTKGQGAEEVGDVVBESGSLPREKLAIQEVPQEAEPYBELAKT
KEVCSGDHTQLTDLSPBEKMLPHPEBIVSEVEMLSQERIKVQSGFLKLFSSG
LKLTKSKQKGRGGGGGDEBEQYOHITESPESADBEQGSASPEPEELICLE
KSPAPQEAEBEGATSDGEKREGITPMASFKMDVPPKKVRPSESDKEEDLKY
KSATSTESTSGMODEVRAVGEORSEPKRYDTSWMLCVSGSKRARKAS
SSDDSGPRTTIGDGHRAEAKDKKADLPASTVBOQAHSSSPREPAGSPSEGEV
STWESFKRLVTPRKSKSKLEBAEDSGAEQLASEIBPBERESWISIKKFTIGRKKR
ADGQEOAVEDSGPEINEDPDVPAVAVLSEYVAVEREKLAEARQADENVLPQK
AVYVEBELSKTLVHTVSAVIDGTRAVTSAEBRSWSMTSASMTPELHAEVATPVG
EYTEKDTAEATPALAQLTGGKADHDIVTSEVPTSAVAETTELAELELTA
SGAERTTMVSAVSOLOSPTTEATTPVQVEGMDLDEBEROTQVLOAVADKVK
EDSOVPAOTTLQAGKALEKVEBESEVALTEKRVNVEGEPQAEETHLAQS
ETVQATPESLEVEVTEVDVRAITTCOVIAHQQLMEQAVAPSSSETLIDSETNGSTPLA

DSPTPGTODDETVDSDNSAIAAVKQSVTEBEAAAQOTESPSSPPAOEHRK
PKRDVLEPTOTALAAGAPILAKAEVQDEBAGCPDDEKVDGCVLELTPVHTGPS
QKADTLRDSVNEVACQETESNEBDSISPEREKWTDVKEETETKTOASBEHBO
ETAPABEGHPRKPVLTADMPHBERGALSLSPBLPODDADCLEVOVSSTPV
TQTEAVKKEEVAIVASEMBDSLECAQSLPAEKISETGYTLQGEEDTVQGPES
QAESIPITVPAPEISLHSDLOREVSASQKORDEDEKPPAGPDAKSEAEKILR
ADPEILESKSKIVOSVITQTVADPARTETAPETHASLOKOVPMQADSGCAQOM
LDKDESLQVSRDQGLTSAVAOGLANSDSEKMSKASMTTLTAVSASVKSEVEL
PLQCKDEKHAADGPGPOHSLAKADADSGNLTKESPDITNGPKLTFEBSDLKEENKQAO
TEEDDLOEPKDLTES"
BASE COUNT 1840 a 1447 c 1779 g 1131 t
ORIGIN
Alignment Scores:
Pred. No.: 7,31e-230 Length: 6197
Score: 6663.00 Matches: 1357
Percent Similarity: 88.90% Conservative: 76
Best Local Similarity: 84.18% Mismatches: 156
Query Match: 82.53% Indels: 23
DB: 10 Gaps: 11
US-09-902-432-4 (1-1596) x AF326228 (1-6197)
QY 1 MetGlyAlaGlySerSerThrGlnGlnArgSerProGlnGlnProAlaGlySerAspThr 20
Db 140 ATGGGTGACGACGATTCACACCGACGACGAGCCCGGACGACCGGCGGAGACGACACG 199
QY 21 ProSerGlnLeuValLeuSerGlyHisGlyProAlaAlaGlnAlaSerGlyValAlaGly 40
Db 200 CCGACGACGCTGAGACTCAGTGGCCATGGCCCGGACGGAAGCGTCCGGAGACACTGGA 259
QY 41 AspProAlaAspAlaAspProAlaThrIysLeuProGlnLysAsnGlyGlnLeuSerSer 60
Db 260 GATCCCGCTGACGCGGACCCGCCACCAAGCTCCACAGAAAGATGTCAGCTGCTCC 319
QY 61 ValAsnGlyValAlaGlnGlnGlyAspValHisValGlnGlnGlnGlnGlnGln 80
Db 320 GTCAATGGTGAAGTCAACAGAGATGTCACGTCACAGAGAAAGCAGAGATGGGCA 379
QY 81 GlnGlnGlnValAlaAspGluAspValGlyGlnArgGlnSerGluAspValArgGlnLys 100
Db 380 GAGGAAGAAGTCACTGTTGAAGATGTTGACAGAGAGTCAAGAAAGATGTAAGAAAAA 439
QY 101 AspArgValGlnGlnMetAlaAlaAsnSerThrAlaValAlaGluAspIleThrLysAspGly 120
Db 440 GACCCAGCTTAAAGAAATGACGACCCAGTTCACAGTTGTTGAAGATATCAAAAGACGAG 499
QY 121 GlnGlnGlnThrSerGlnIleIleGlnGlnIleProAlaSerGluLysAsnValGlnGln 140
Db 500 CAGGAGAAACACCGGAATTAATCAACAGATCCCTGCTTCAGAGCAATGTGAGAA 559
QY 141 MetValGlnProAlaGlnSerGlnAlaAsnAspValGlyPheLysLysValPheLysPhe 160
Db 560 ATGGCGACGCGTGTGAGTCCCAAGCTAATGACGTCCGCTTCAAGAAAGTATTAAATT 619
QY 161 ValGlyPheLysPheThrValIysLysAspLysAsnGlnLysSerAspThrValGlnLeu 180
Db 620 GTTGTGTTTAAATTACCGTGAAGAAAGATTAAGAAAGTCAAGATCCGTCAAGCTA 679
QY 181 LeuThrValIysLysAspGlnGlyGlnGlnGlnAlaGlnIleAspValGlyValGlyAspHis 200
Db 680 CTCACGTGTCAGAAAGATGAGAGGCGCAAGGCGCAAGGCTCCGTGAGACGAGAACAC 739
QY 201 GlnGlnProSerValGlnThrAlaValGlyGlnSerAlaSerLysGlnSerGlnLeuLys 220
Db 740 CAAAGGCCCGAGTGAAGACC---GTCCGCGAATCAGGATCCAAAGAAAGTACGTGAAG 796
QY 221 GlnSerThrGlnLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 240
Db 797 CAATCCACAGAAACCAAGAGGACCCCTGAAGCAAGACAGACGACGACGAAATTTCC 856
QY 241 LeuGlnAlaGlnSerAspGlnAlaAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 260

Db 857 CTTCAAGCCGAATCTGGTCAAGGACCGAGGAAGACGCCAAAGATGGAGAAGAAAC 916
Qy 261 GlnGluLysGluProThrLysSerProGluSerProSerSerProValAenSerGluThr 280
Db 917 CGAGAGAAAGAACCTACCAAGCCCTAGAACTCTCCGACCGAGCCCTGTTCAGCAATGAGACA 976
Qy 281 ThrSerSerPheLysLysPheThrHisGlyTrpAlaGlyTrpArgLysLysThrSer 300
Db 977 ACATCTTCTTCAAGAAATTTCTTCACTCAGGGTGGCCGTGGCCCAAGAGACCAGC 1036
Qy 301 PheLysLysSerLysGluAspAspLeuGluThrAlaGluLysArgLysGluGlnGluAla 320
Db 1037 TTCAAGAAACCAAGGAAGATGATCTGAAACTTCCGAGGAAGAAAGAGCAAGAGGCT 1096
Qy 321 GluLysValAspGluGluLysGluLysThrGluProAla---SerGluGluGlnGlu 339
Db 1097 GAAAAAGTAGACGAGGAAGAGGGGAAAGACAGAGCCAGCCAGCGAGGAGGAG 1156
Qy 340 ProAlaGluAspThrAspGlnAlaArgLeuSerAlaAspTyrGluLysValGluLeuPro 359
Db 1157 CTTGCAAGGCGACAGACCGAGCGAGGTGTGACGCCACTATGAGAAAGTGGAGTTGCCT 1216
Qy 360 LeuGluAspGlnValGlyAspLeuGluAlaSerSerGluGluLysCysAlaProLeuAla 379
Db 1217 TTGGAAGACCGAGTGGTGACCTGGAGGCATTGTGCG---GAGAAAGTGTGCTCTTTGGCA 1273
Qy 380 ThrGluValPheAspGluLysMetGluAlaHisGlnGluValValAlaGluValHisVal 399
Db 1274 ACGGAAGTGTGTGATGAGAAGACGGAAGCCACCAAGAAAGTGTGTCAGAGGTCACAGTG 1333
Qy 400 SerThrValGluLysThrGluGluGlnGlnGlyGlyGlyGluAlaGluGlyGlyVal 419
Db 1334 AGCACCGTGGAGAAGATGACGAAGAGGGCAA---GGAGGACAGAGGTGGAAAGGGGATGTG 1390
Qy 420 ValValGluGlyThrGlyGluSerLeuProProGluLysLeuAlaGluProGlnGluVal 439
Db 1391 GTGGTGAAGATCGGAGAAATCTTCTCCCTGAGAAATCTGGCTGAGACCCAGGAGTTC 1450
Qy 440 ProGlnGluAlaGluProAlaGluGluLeuMetLysSerArgGluMetCysValSerGly 459
Db 1451 CCCAGGAAGCTGAGCTGTGGAGGAGCTGATGAAGACCAAGAAAGTATGCGTCTCTGGG 1510
Qy 460 GlyAspHisThrGlnLeuThrAspLeuSerProGluGluLysThrLeuProLysHisPro 479
Db 1511 GGTGACCATACTAGCTGACAGATCTAAGTCTTGAAGAGAGATGCTTACCCAAACACCCC 1570
Qy 480 GluGlyIleValSerGluValGluMetLeuSerSerGlnGluArgIleLysValGlnGly 499
Db 1571 GAAGGCATTTGTCAGTGAGTGGAGATGCTGCTCTCAGAGAGAAATCAAGGTACAGGGA 1630
Qy 500 SerProLeuLysLysLeuPheSerSerSerGlyLeuLysLysLeuSerGlyLysLysGln 519
Db 1631 AGTCCCTGAAGAAGCTCTTTCAGCAGTTCGGGCTTAAAGAAGCTCTCCGGGAAGAGCAG 1690
Qy 520 LysGlyLysArg---GlyGlyGlyAspGluProGlyGluTyrGlnHisIleHis 538
Db 1691 AAGGGGAAGAGAGAGGAGGCGGGGAGATCAAGAGCCAGGAGAAATACCAACACATTCAA 1750
Qy 539 ThrGluSerProGluSerAlaAspGluGlnLysGlyLeuSerSerAlaSerProGlu 558
Db 1751 ACCGAGTCCCCAGAGTGTCTGACGACGAAGGAGAGAGTCTGCTCTTCCCCCTGAA 1810
Qy 559 GluProGluGluThrThrCysLeuGluLysGlyProLeuGluAlaProGlnAspGlyGlu 578
Db 1811 GAGCCCGAGGAGATCGCGTGTCTGGAGAAGGGGCCATCGGAAGCACCCCAAGAGCGGAA 1870
Qy 579 AlaGluGluGlyThrThrSerAspGlyGluLysLysArgGluGlyIleThrProThrAla 598
Db 1871 GCTGAGGAGGAGGAGCTTCCGACGGAGAGAAGAAAGGAGGAGATCACCCCTGGGCA 1930
Qy 599 SerPheLysLysMetValThrProLysLysArgValArgArgProSerGluSerAspLys 618
Db 1931 TCCTTTCAAAAGATGGTGAACACCAAGAAACCGGTCCGAAGACCTTCTTGAGAGCGCAAG 1990

Qy 619 GluGluGluLeuGluLysValLysSerAlaThrLeuSerSerThrAspSerThrValSer 638
Db 1991 GAAGAAGAGCTGGATTAAGGTCAAGAGTCCACCTTGTCTCCACGAGAGCACGGCTCT 2050
Qy 639 GluMetGluAspGluValLysThrValGlyGluGluGlnLysProGluGluProLysArg 658
Db 2051 GGATTCAGATGAGTCAAGCGGTGGCGAGGAGCAAGGTCAAGAGAGCCAAAGCGC 2110
Qy 659 ArgValAspThrSerValSerTrpGluAlaLeuIleCysValGlySerSerLysLysArg 678
Db 2111 AGSGTGGATACTTCAGTGTCTTGGGAGGCGTTGATTTGTTCGATCGTCCAAGAGAGA 2170
Qy 679 AlaArgLysAlaSerSerSerSerAspGluGlyGlyProArgThrLeuGlyGlyAspSer 698
Db 2171 GCAGAGAAAGCATCTCTTCAGATGATGAAGGAGGCGCAAGACACTGGGAGGGATGGC 2230
Qy 699 HisArgAlaGluGluAlaSerLysAspLysGluAlaGlyThrAspAlaValProAlaSer 718
Db 2231 CACAGAGCGGAGGAGGTAGCAAAAGCAAAAGACA-----GATGCTCTTCTGCCAGC 2284
Qy 719 ThrGlnGluGlnAspGlnAlaGlnGlySerSerSerProGluProAlaGlySerProSer 738
Db 2285 ACCCAGACACAGACCAAGCGCACGGAAGTTCCTCACCCGAGCCAGCTGGAGCCCTTCT 2344
Qy 739 GluGlyGluGlyValSerThrTrpGluSerPheLysArgLeuValThrProArgLysLys 758
Db 2345 GAAGGGAGGGCGCTCTCCACCTGGGAGTCATTTAAGAGATTAGTACTCCACGAAAAAAA 2404
Qy 759 SerLysSerLysLeuGluLysAlaGluAspSerSerValGluGlnLeuSerThrGlu 778
Db 2405 TCCAAGTCAAACTGGAAGAGAGAGCCGGAAGACTCCGGTGCAGAGCAGTTGGCCTCCGAG 2464
Qy 779 IleGluProSerArgGluGluSerTrpValSerIleLysLysPheIleProGlyArgArg 798
Db 2465 ATCGAACCAAGTAGAGAGGAATCTTGGGTTTCCATTAAAGAAATTATTCTTGGACGGCG 2524
Qy 799 LysLysArgAlaAspGlyLysGlnGluGlnAlaThrValGluAspSerGlyProValGlu 818
Db 2525 AAGAAAGGCGCAGATGGGAAGCAAGAACAGGCGCGCTTGAAGACTCGGGGCCAGGAG 2584
Qy 819 IleAsnGluAspAspProAsnValProAlaValValProLeuSerGluTyrAsnAlaVal 838
Db 2585 ATCAATGAGGAGCAGACCCCGACGCTCCAGCTGTGTGCTCTGTCTGAGTACGATCGGTA 2644
Qy 839 GluArgGluLysMetGlu-----AlaGlnGlyAsnThrGluLeuProGlnLeuLeu 855
Db 2645 GAGAGAGAGAGCTGGAAGCGCAGCGAGCTCAGGAGAACCTGGAGCTGCCACGCTGAAG 2704
Qy 856 GlyAlaValTyrValSerGluGluSerLysThrLeuValHisThrValSerValAla 875
Db 2705 GGGGCTGTGTATGTCTCAGGAGCTTAGTAAGACTCTCGTTTCCACTGTGAGTGTGCG 2764
Qy 876 ValIleAspGlyThrArgAlaValThrSerValGluArgSerProSerTrpIleSer 895
Db 2765 GTCAATTGATGGACCGAGGCGAGTCAACAGTCGGAAGAGCGGTCCCTTCGTGGATATCT 2824
Qy 896 AlaSerValThrGluProLeuGluHisThrAlaGlyGluAlaMetProProValGluGlu 915
Db 2825 GCTTCATCAGCAGAACTCTTGAGCAGCGAGGAGGAGTGCCACCGCTGTGTGAGAG 2884
Qy 916 ValThrGluLysAspIleAlaGluGluThrProValLeuThrGlnThrLeuProGlu 935
Db 2885 GTCACTGAAAAGACATCACTCAGAGAGCAACTCTCGTCACTCGCCAGACTTTACGAGG 2944
Qy 936 GlyLysAspAlaHisAspAspMetValThrSerGluValAspPheThrSerGluAlaVal 955
Db 2945 GGCAGAGATGCCCATGACGACATAGTCAACAGTGAAGTGGATTTTACCTCAGAAAGCAGTG 3004
Qy 956 ThrAlaThrGluThrSerGluAlaLeuArgThrGluGluValThrGluAlaSerGlyAla 975
Db 3005 ACAGCGCAGAAACACACAGAGGCGCTCCGCGCTGAAGAACTTACCGAAGCATCAGGGCA 3064


```

MEDLINE      21322683
PUBMED       11429284
REFERENCE    2 (bases 1 to 5902)
AUTHORS      Mesbah, K., Babinet, C. and Barra, J.
TITLE        Direct Submission
JOURNAL       Submitted (06-DEC-2000) Immunologie, Institut Pasteur, 25 rue du
              Docteur Roux, Paris 75015, France
FEATURES
source
  1. 5902
  Location/Qualifiers
    /organism="Mus musculus"
    /db_xref="taxon:10090"
    /chromosome="10"
    /tissue type="testis"
    150..4899
    /codon_start=1
    /product="testis-specific SSeCKs"
    /protein_id="AAK16152.1"
    /db_xref="GI:13236408"
    /translation="MAASSTVVEDITDKDQEBTEPIIEOI PASESNVREMAOAAESQAA
    NDVGKKVFKFYGFKFTVKKDKNEKSDTVQLITVKKDEGEAEASVGADGHPGPGVET
    VGESAKSEKSLQSTKQEGTLKQASSTETPLQAESGGGTGTEEAADKQGEENREKEPT
    KPLESTPSVNETTSSFKFFTHGWAGWRKKTSTFKPKEDDLLETSEKKEQEAKEKVD
    EBEGETEPAPAEQEPARECTQDARLSADYKEVLPLEDQVGDLSALSEKCAPLATEV
    FDEKTEAHQEVVAEHVSTVKMTKGGAEGEVGVVVEGSGESLPPEKLABTQBPVQ
    EAPVEELMKTKEVCSGGDHTQDLSPEEKMLPKHPGIVSEVEMLSQSERIKVQG
    SPLKLFSSGLKLSGKKQKRGCGGDDPEPGYQHIOQTESPESADQKGESSASS
    PEPEPIACLEGPSEAPQEAABEGATSDGKKREGITPWAISFKKMTVPKKVRFRPS
    ESDKEELDKVKSATLSTESTASGMQDEVRVAGEEQRSEEPKRRVDTQSVSWEALICV
    GSKKRRARASSDDGGPRTGLGGHRAEASKQKEADALPASTQEOQDQAHGSSSPE
    PAGESPBGVSTWFSFKRLVTPRKKSLEKREARSDGAEOLESEIEPRESFVSI
    KKFIPGRKRKRDAGQEOQAAVEDSGPGEINLEDPPDPAVVPVLPSEVDAREKLEQRA
    QNVELPQLKGVAVSEELSKTLVHTVSAVIDGNRAVTSABERSPSWISASMTPLE
    HAEGVATPVGEETKIDTAATPALAQLPGCKDAHDIVTSEVDFTSVATAAEIT
    EALRAEELTEASGAETTDTEATPQVEEGGMLDTEQERQTO
    AVLQAVADKQSTVOATPOTLQRAQKALEKVEVEEDSEVLATEKEKDVVPEPGLT
    BAETHLAQGSTVOATPESLPEVETEDVDRATTCOVTKHOOLMEQAVAPSSSTLT
    DSEWTGSTPLANDSDTPNGTQOQDETVDSDSNALAAVQSOVTEEEAAAQATSGPSTPS
    SPAPQEHKRGDRDVFETQALAGAVFPLAKAEVQGEAGQFDGKVKDQGCCKE
    LSPVHTGPSNOKTADLRDSEVMEVARCOETSENEEOSISPEKREMGTDVBEKETET
    KTEQASDEEHQTAAPAEHGHGTHPKPVLTDMPHSERKALGSLGSPSLPDQDKADCI
    EVQVQSDPTPTVTQTAATKVEETVATSEDSLECAQAQSLPAEKLSSETGGYGLQH
    GEDTVPQGPESQAESPIIIVTPAPESILHSDLOREVSASQKORSDEDNKPDAGPDAAG
    KESAREKILRAEPLEILESKNKIVSQIOTAVDQFARTETATETHASDMLQNOQPVV
    MQADSQGQOQMLKDESDLIQVSPQDGLTSAVQOGLVAVSDSEGMSKASEMITTLAVE
    SASVSEMKNAQTEEDDLQEPKGLDLES"
    DALKSEMNKAQTEEDDLQEPKGLDLES"
BASE COUNT   1774 a 1351 c 1677 g 1100 t
ORIGIN
Alignment Scores:
Pred. No.:    2,61e-215      Length:    5902
Score:         6255.00       Matches:   1276
Percent Similarity: 88.58%   Conservative: 73
Best Local Similarity: 83.78% Mismatches:   151
Query Match:    77.48%      Indels:     23
DB:             10          Gaps:       11

US-09-902-432-4 (1-1596) x AF326230 (1-5902)

Qy  90  GlyGlnArgGluSerGluAspValArgGluLysAspArgValGluGluMetAlaAlaAsn 109
    |||
Db  112 GGACAGAGAGGTGCAAGATGTGAAGAGAAAAGACCGAGCTTAAGAAAATGGCAGCCAGT 171
    |||

Qy  110 SerThrAlaValGluAspIleThrLysAspGlyGlnGluThrSerGluIleIleGlu 129
    |||
Db  172 TCACAGTGTGTTAGATATATCAAAAGGACGACGAGAGAAAACCGGNAATATCGAA 231
    |||

Qy  130 GlnIleProAlaSerGluAsnValGluGluMetValGlnProAlaGluSerGlnAla 149
    |||
Db  232 CAGATCCCTGCTTCAGAGACCAATGTGGAAGAAAATGGCGCAGCGTCTGCTGAGTCCCAAGCT 291
    |||

Qy  150 AsnAspValGlyPheLysLysValPheLysPheValGlyPheLysPheThrValLysLys 169
    |||

```

```

Db  292 AATGAGCTCGGCTTCAAGAAGGTATTAAATTTGTTGGTTTTAAATTCACGCTGAAGAAG 351
Qy  170 AspLysAsnGluLysSerAspThrValGlnLeuLeuThrValLysLysAspGluGlyGlu 189
    |||
Db  352 GATAAAAACGAAAAGTCAGATACCGTCCAGCTACTCCTGTCAAGAAGGATGAAGCGAA 411
    |||

Qy  190 GlyAlaGluAlaSerValGlyAlaGlyAspHisGlnGluProSerValGluThrAlaVal 209
    |||
Db  412 GGGGCGAAGAGCTCGCTCGGAGCAGGAGACCACCAAGAGCCCGGAGTGAGAGACC---GTC 468
    |||

Qy  210 GlyGluSerAlaSerLysGluSerGluLeuLysGlnSerThrGluLysGlnGluGlyThr 229
    |||
Db  469 GCGCATCAGCATCCAAAGAAAGTGAGCTGAAGCAATCCACAGAGAGCAAGAGGCACC 528
    |||

Qy  230 LeuLysGlnGluGlnSerThrGluIleProLeuGlnAlaGluSerAspGlnAlaAla 249
    |||
Db  529 CTGAAGCAAGCAGCAGCAGCAGCAAGAAATTCCTCTTCAAGCCGAAATCTGGTCAAGGGACC 588
    |||

Qy  250 GluGluGluAlaLysAspGluGlyGluGluLysGlnGluLysGluProThrLysSerPro 269
    |||
Db  589 GAGGAAGAAGCAGCCAAAGATGGAGAGAAAACCCGAGAGAAAACCTTACCAAGCCCTTA 648
    |||

Qy  270 GluSerProSerSerProValAsnSerGluThrThrSerSerPheLysLysPhePheThr 289
    |||
Db  649 GAATCTCCGACAGCCCTGTGACCAATGAGACAACATCTTCTTCAAGAAATTTCTTCACT 708
    |||

Qy  290 HisGlyTrpAlaGlyTrpArgLysLysThrSerPheLysLysSerLysGluAspAspLeu 309
    |||
Db  709 CACGGGTGGCGCGCTGGCGCAAGAGACCCAGCTTCAAGAAAACCAAGAGGAGATGATCTG 768
    |||

Qy  310 GluThrAlaGluLysArgLysGluGlnGluAlaGluLysValAspGluGluLysGlu 329
    |||
Db  769 GAAATCTCCGAGAAGAGAAAGGAGCAAGAGGCTGAAAAGATAGACGAGGAAGAGGGGAA 828
    |||

Qy  330 LysThrGluProAla---SerGluGluGlnGluProAlaGluAspThrAspGlnAlaAa 348
    |||
Db  829 AAGACAGAGCCAGCCAGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 888
    |||

Qy  349 LeuSerAlaAspTyrGluLysValGluLeuProLeuGluAspGlnValGlyAspLeuGlu 368
    |||
Db  889 TTGTCAAGCGACTATGAGAAGGTGGAGTTGCTTGGGAAGACAGGTCGTGACCTGGAG 948
    |||

Qy  369 AlaSerSerGluGluLysCysAlaProLeuAlaThrGluValPheAspGluLysMetGlu 388
    |||
Db  949 GCATTGTGCG---GAGAAGTGTGCTCTTTGGCAACGGAAGTGTGTTGATCAGAAGACGGAA 1005
    |||

Qy  389 AlaHisGlnGluValValAlaGluValHisValSerThrValGluLysThrGluGluGlu 408
    |||
Db  1006 GCCCAACCAAGAGTGTGTCAGAGGTCCACGTGAGCAGCCGTGGAGAAATGACGAAAGGG 1065
    |||

Qy  409 GlnGlyGlyGlyGlyAlaGluGlyGlyValValValValGluGlyThrGlyGluSerLeu 428
    |||
Db  1066 CAA---GGAGGAGCAGAGGTGGAGGGGATGTGTTGGTGGAGAGGATCGGAGAAATCTT 1122
    |||

Qy  429 ProProGluLysLeuAlaGluProGlnGluValProGlnGluAlaGluProAlaGluGlu 448
    |||
Db  1123 CCCCTTGAAAACCTGGCTGAGACCCAGGAGGTCCCCCGAGAAAGCTGAGCCTGTGGAGGAG 1182
    |||

Qy  449 LeuMetLysSerArgGluMetCysValSerGlyGlyAspHisThrGlnLeuThrAspLeu 468
    |||
Db  1183 CTGATGAAGCAAAAGATGTCGCTCTCTGGGGGTGACCATACTCAGCTGACAGATCTTA 1242
    |||

Qy  469 SerProGluGluLysThrLeuProLysHisProGluGlyIleValSerGluValGluMet 488
    |||
Db  1243 AGTCTGAAGAGAGATGCTACCCAAACACCCCGAAGGCAATGTCAGTGAGGTGGAGATG 1302
    |||

Qy  489 LeuSerSerGlnGluArgIleLysValGlnGlySerProLeuLysLysLeuPheSerSer 508
    |||
Db  1303 CTGTCTCTCTCAGAGAGAAATCAAGGTACAGGGAAGTCCCTTCAAGAAGCTCTTTCAGCA 1362
    |||

Qy  509 SerGlyLeuLysLysLeuSerGlyLysLysGlnLysGlyLysArg---GlyGlyGlyGly 527
    |||
Db  1363 TCGGGCTTAAAGAAGCTCTCCGGGAAGAGAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAG 1422
    |||

```

QY 528 AspgIuGlupProgluYgluYrGlnHisIleHsThrGluSerProgluSerAlaAspGlu 547
 Db 1423 GATGGAAGACCCAGAGAAATACAAACATTCAACCGAGTCCCAAGAGTGTGACGAG 1482
 QY 548 GlnuYsgIuGlusSerAlaSerSerProgluGluProgluGluThrCysLeuGlu 567
 Db 1483 CAGAAAGGAGAGACTCTGCTCTCTCTCTGAAAGAGCCGAGAGATGCCGTCTCGAG 1542
 QY 568 LysglYProleuGluAlaProgluAspGlygluAlaGluGluGluYrThrSerAspGly 587
 Db 1543 AAGGGGCGCATCGAAGACACCCAGAAAGCGAAGCTGAGAAAGAGCGACTTCCGACGGA 1602
 QY 588 GlnuYslyAspArgIuGluYrIleThrProTTPAlaSerPheLyslyMetValThrProlys 607
 Db 1603 GAGAAAGAAAAGGAGAGATCAACCCCTGGGCACTCTTCAAAAAGTGTGTACACCCAAAG 1662
 QY 608 LysArgValArgArpProSerGluSerAspLysGluGluGluLeuGluYrValLysSer 627
 Db 1663 AAACGGGTCCGAAGACTTCTGAGAGCGACAGAAAGAAAGCTGGATTAAGTCAAGAGT 1722
 QY 628 AlaThrLeuSerSerThrAspSerThrValSerGluMetGlnAspGluValYrThrVal 647
 Db 1723 GCCACCTTGTCTCCACGAGAGACAGGCGTCTGAATGACAGATGAGGTCAAGCGGTT 1782
 QY 648 GlyGluGluGluYrProgluGluProgluYrAspArgValAspThrSerValSerTpglu 667
 Db 1783 GCGGAGAGACAAAGCTCAAGAGAGCGAAAGCGAGGTGATCTTCACTGTCTTGGAG 1842
 QY 668 AlaLeuIleCysValGlySerSerLysArgAlaArgLysAlaSerSerSerAspAsp 687
 Db 1843 GCGTGTATTTGTGCGATGCTCCAGAAAGAGCGAGAGAGCATCTCTTCAAGTAT 1902
 QY 688 GluGlyGlyProArgThrLeuGlyYrAspSerHisArgAlaGluGluAlaSerLysAsp 707
 Db 1903 GAAGAGGGCCAAAGACACTGGAGGGGATGGCCACAGAGCGAGAGGTAGCAAAAGC 1962
 QY 708 LysGluAlaGlyThrAspAlaValProAlaSerThrGlnGluGlnAspGlnAlaGlnGly 727
 Db 1963 AAGAAGACA-----GATGCTCTTCTGCGACACCCAGAAACAAACCTGGAGAGAGGCCA 2016
 QY 728 SerSerSerProgluProAlaGlySerProSerGluGluGluYrValSerThrTpglu 747
 Db 2017 AGTTCTCTACCCAGAGCAAGTGAAGCCCTTCTGAAGGGAGGGCGTCTTCCACTGGAG 2076
 QY 748 SerPheLysArgLeuValThrProArgLysLysSerLysSerLysLeuGluGluYrVal 767
 Db 2077 TCATTTAAGAGATTACTCACTCCACGAAAAAATCCAAAGTCAAACTGGAGAGAGGCC 2136
 QY 768 GluAspSerSerValGluGluLeuSerThrGluIleGluProSerArgIuGluGluSerTTP 787
 Db 2137 GAAGACTCCGCTGCAAGACAGTTGGCTCCGAGATCGAACCAGTGAAGAGGAGATTTGG 2196
 QY 788 ValSerIleLysLysPheIleProArgArgLysLysArgAlaAspGlyLysGlnGlu 807
 Db 2197 GTTTCATTAAAGAAATTTATCTTGACGGCGGAAAGAAAGGCAATGGGAAGCAAGAA 2256
 QY 808 GlnAlaThrValGluAspSerGlyProValGluIleAsnGluAspAspProAsnValPro 827
 Db 2257 CAGGCCCGCTTGAACACTCGGGGCCAGAGAGATCAATAGGACACACCCCGACGTCCA 2316
 QY 828 AlaValValProLeuSerGluYrAsnAlaValGluArgGluLysMetGlu----- 844
 Db 2317 GCGTGTGCGCTGTCTGTGATCGATGCGGTAGAGAGAGAGAGAGTGGAGCGCAGCGA 2376
 QY 845 AlaGlnGlyAsnThrGluLeuProGlnIleLeuGlyAlaValYrValSerGluGluLeu 864
 Db 2377 GCTCAGAGAGAACTGAGCGTCCAGCTGAAGGGGGCTGTGATGTGTCTGGAGGCTT 2436
 QY 865 SerLysThrLeuValHisThrValSerValAlaValIleAspGlyThrArgAlaValThr 884
 Db 2437 AGTAAAGCTCTGTTCACTGTAGTGTCTGCGGTCAITGATGGAGCCAGGAGAGTCAAC 2496

QY 885 SerValGluArgSerProSerProSerTrpIleSerAlaSerValThrGluProLeuGluHis 904
 Db 2497 AGTCCGAAGAGCGGTCCCTTCTGTGATATCTGCTTCAAGACAGAACTCTTGAGAC 2556
 QY 905 ThrAlaGlyGluAlaMetProProValGluGluValThrGluYrAspIleIleAlaGlu 924
 Db 2557 GCAGAGGAGTGGCCACACCCCTGTGGAGAGGTCACTGAAAAAGACATCACTGCAGAA 2616
 QY 925 GlnThrProValLeuThrGlnThrLeuProgluGluYrAspAlaHisAspAspMetVal 944
 Db 2617 GCACTCTCTGCACTCGCCCAACTTACAGGGGGCAAAAGATGCCCAAGAGACATAGTTC 2676
 QY 945 ThrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeu 964
 Db 2677 ACCAGTAGTGATTTTAACTCAGAGACATGACGCCGAGAAACCAAGAGCGCTTC 2736
 QY 965 ArgThrGluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSerAla 984
 Db 2737 CGCGTGAAGAACTTACCGAAGCATCAGGGGCAAGAAAGACACAGACATGTGTCTGCA 2796
 QY 985 ValSerGluLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGluVal 1004
 Db 2797 GTTTCACAGCTGTCCGACTCCCGGACACAGAGAGACCCCGCTCAGAGAGTA 2856
 QY 1005 GluSerGlyValLeuAspThrGluGluGluYrGlnThrGlnAlaIleLeuGlnAla 1024
 Db 2857 GAGGTGCGATGCTAGATACGAGAAACAGAGAGCGCCAGACAGCGCGTCTCCAAACC 2916
 QY 1025 ValAlaAspLysValLysGluGluSerGlnValProAlaThrGlnThrValGlnArgThr 1044
 Db 2917 GTTGACAGCAAAAGTAAAGAGAGACTCCAGGTGCTGCAACCCAGACTCTGACAGAGCA 2976
 QY 1045 GlySerLysAlaLeuGluYrValGluGluValGluGluAspSerGluValLeuAlaSer 1064
 Db 2977 GGGCCGAAGACACTGAGAAAGTGAAGAGGTAGAGGAGACTCCGAGAGTCTGTGCTTCC 3036
 QY 1065 GlnuYsgIuYrAspValMetProLysGlyProValGlnGluAlaGlyValGluHisIleu 1084
 Db 3037 GAGAAAGAGAAAGATTTGTGCGGAGAGACCGGTGCGAAAGCTGAAACTGAGCATCTT 3096
 QY 1085 AlaGlnGlySerGluThrGlyGlnAlaThrProGluSerLeuGluValProgluValThr 1104
 Db 3097 GCACAGGGCTCCGAAACACTGACAGCTACCCACAGAGACCTTGAAGTCTCGAAGTCA 3156
 QY 1105 AlaAspValAspHisValAlaThrCysGlnValIleLysLeuGlnGlnLeuMetGluGln 1124
 Db 3157 GAGGATGTAGACCTGCGCACACATGCGAGGTATCAAGACACCAAGCACTGATGAGCAG 3216
 QY 1125 AlaValAlaProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrPro 1144
 Db 3217 GCTGTGGCCCTGAGTCAATCTGAACCTTGACAGACAGTGAACAAATGAGAAATCTCCC 3276
 QY 1145 LeuAlaAspSerAspThrAlaAspGlyThrGlnGlnAspGluThrIleAspSerGlnAsp 1164
 Db 3277 CTCGCAAGTTCAGACACTCCAAACCGGACACAGCAAGACGAGACCGTTGACAGCCAGAC 3336
 QY 1165 SerLysAlaThrAlaAlaValArgGlnSerGlnValThrGluGluGluAlaAlaThrAla 1184
 Db 3337 AGTAATGCCATATGCCCGCTCAAGCAGTCAAGCTCAAGAAAGAGAGGAGCTCTGCT 3396
 QY 1185 GlnuYsgIuGluProSerThrLeuProAsnAsnValProAlaGlnGluGluHisGlyGlu 1204
 Db 3397 CAGACGGAGGGGCTTCAACA---CCATCTAGTTTTCAGGCCAGAAAGAACAGGGGA 3453
 QY 1205 GluProGlyArgAspValLeuGluProThrGlnGlnGluLeuThrAlaAlaAlaValPro 1224
 Db 3454 AAACCAAGAGAGGATCTTCAAGACCA---CAAGCGCTGTGCTCGGGGCAATGCTCT 3510
 QY 1225 ValLeuAlaLysThrGluValGlyGlnGluGluValAspTrpLeuAspGlyGluYr 1244
 Db 3511 ATTCTGCAAAAGCTGAGGTGGTCAAGAGGGGTGAGCGTGGCACTTGTATGAGAAAAA 3570
 QY 1245 ValLysGlu-----GluGlnGluValPheValHisSerGlyProAsnSer 1259

```
Db 3571 GTCAAAGCGACAGTGTGTTAAAGAACTGGAGTGCTGTGCACACTGGACCCCAACAGT 3630
Qy 1260 GlnlyslAlaAAspValThrTyrAspSerGluValMetGlyValAlaGlyCysGlnGlu 1279
Db 3631 CAAAAGACTGCTGACTTGACACGTCAGAGTGAAGTAAATGGAAGTGGCCAGATGTCAGGAA 3690
Qy 1280 LysGluSerThrGluValGlnSerLeuSerLeuGluGluGlyGluMetGluThrAspVal 1299
Db 3691 ACTGAGAGTAAATGAAGAACACAGAGTATTAGCCCGAGAAAGAGAGAGTGGGAACCGAGTT 3750
Qy 1300 GluLysGluLysArgGluThrLysProGluGlnValSerGluGluGlyGluGlnGluThr 1319
Db 3751 GAAAGAGGAGAAACAGAGACCCACAGACAGACCCAGTGAAGAACATGACAGCAAAACA 3810
Qy 1320 AlaAlaProGluHisGluGlyThrTyrGlyLysProValLeuThrLeuAspMetProSer 1339
Db 3811 GCTGCTCTGAGCATGAAGAAACCAACCCCTAAGCCAGTCTCTGACAGCTGACATGCCTCAC 3870
Qy 1340 SerGluArgGlyLysAlaLeuGlySerLeuGlyCysProSerLeuProAspGlnAsp 1359
Db 3871 TCAGAGAGGGGAAAGGCACTGGCGAGCCTTGAAGGAAAGCCCTTCTCTCCACAGACCAAGAC 3930
Qy 1360 LysAlaGlyCysIleGluValGlnValGlnSerLeuAspThrThrValThrGlnThrAla 1379
Db 3931 AAACAGATTGCATAGAGTTCAAGTTCAAGTTCAGACACACCCAGTCACTCAAAACCAACC 3990
Qy 1380 GluAlaValGluLysValIleGluThrValValIleSerGluThrGlyGluSerProGlu 1399
Db 3991 GAAGCTGTGAAAAAGGTCGAAGAAACTGTGCAACTTCAGAGATGGATGAAGTTTGGAG 4050
Qy 1400 CysValGlyAlaHisLeuLeuProAlaGluLysSerSerAlaThrGlyGlyHisTrpThr 1419
Db 4051 TGTCCAGGTGCGCAATCATTTACCAAGCTGAGAAGCTCTCCGAAACCGGTGGCTACGGGACT 4110
Qy 1420 LeuGlnHisAlaGluAspThrValProLeuGlyProGluSerGlnAlaGluSerIlePro 1439
Db 4111 CTTCAGCATGGAGAGACACCGTCCCGAGGGCTGAGTCTCAGGCAGAGTCCATCCCA 4170
Qy 1440 IleIleValThrProAlaProGluSerThrLeuHisProAspLeuGlnGlyGluIleSer 1459
Db 4171 ATAATAGTAACCTCTGCTCTGCTGAAAGCATCTACATTTGACCTTCAAGAGAGAGTGAGC 4230
Qy 1460 AlaSerGlnArgGluArgSerGluGluAspLysProAlaGlyProAspAlaAsp 1479
Db 4231 GCATCCCAAGAACACAGAGATCAGATGAAGATAACAAGCCAGATGCTGTCTGTATGCTGCC 4290
Qy 1480 GlyLysGluSerThrAlaIleGluLysValLeuLysAlaGluProGluIleLeuGluLeu 1499
Db 4291 GGCAGAGAGAGTGAGCAAGAGAGAAAATCCTCAGGGCTGAACCTGAGATCTTGGAACTT 4350
Qy 1500 GluSerLysSerAsnLysIleValLeuAsnValIleGlnThrAlaValAspGlnPheAla 1519
Db 4351 GAGAGTAAGAGCATATAGATTGTCAGAGTGTATCAGACAGCCGTCGACCATGTTGCA 4410
Qy 1520 ArgThrGluThrAlaProGluThrHisAlaTyrAspSerGlnThrGlnValProAla-Cy 1539
Db 4411 CGTACAGAAACAGCCCCCGAAACCCAGCTTCTGATTTACAGATTCAGGTTCTCTGTATG 4470
Qy 1539 sArgLeuAspSerArgGluProAsnArgCysTrpThrLysMetLys----- 1554
Db 4471 CAGGCTGACAGCCAGGGAGGCACAACAGATGCTGTGACAAAGATGAAGCGACCTTCAAGTC 4530
Qy 1555 -AspAlaLysMetLysHisProValProGlnProArgGluAspLeuValLeuThrVal 1574
Db 4531 TCCCCCAAGATGAAACACTCAGTCCGTCAGCCAGGAGGACTTGGGGTTCTGTAGT 4590
Qy 1574 LeuGluAlaTrpAlaGlnProArgLysCysLeuProArgLeuGlnLeuLysAlaProVa 1594
Db 4591 TCTGAAGGCATGACAAAGGCTTCAAGAAATGATCACCAGCTTGCAAGTTGAAGTGCCAGT 4650
Qy 1594 lSerLys 1596
|||||
```

```
Db 4651 GTCAAAG 4657
RESULT 7
LOCUS AR071282 5134 bp DNA linear PAT 18-FEB-2000
DEFINITION Sequence 1 from patent US 5910442.
ACCESSION AR071282
VERSION AR071282.1 GI:7222170
KEYWORDS Unknwn.
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 5134)
TITLE Gelman, I.H.
JOURNAL Tumor suppressor gene
FEATURES Patent: US 5910442-A 1 08-JUN-1999;
LOCATION/Qualifiers
source 1..5134
/organism="unknown"
BASE COUNT 1511 a 1167 c 1451 g 1005 t
ORIGIN
Alignment Scores:
Pred. No.: 1 95e-206 Length: 5134
Score: 6005.00 Matches: 1232
Percent Similarity: 97.33% Conservative: 7
Best Local Similarity: 96.78% Mismatches: 27
Query Match: 74.38% Indels: 14
DB: 6 Gaps: 4
US-09-902-432-4 (1-1596) x AR071282 (1-5134)
Qy 329 GluLysThrGluProAlaSerGluGluGlnGluProAlaGluAspThrAspGlnAlaArg 348
Db 2 GAAAAGACAGACGACGCTCGAGGAGCAGAGAGCCGCGACAGACACAGACACGACCCAGG 61
Qy 349 LeuSerAlaAspTyrGluLysValGluLeuProLeuGluAspGlnValGlyAspLeuGlu 368
Db 62 TTGTGACGAGACTACGAGAGGTGGAGTGCCTTTTGGAAAGCAGAGTGTGTGACCTGGAG 121
Qy 369 AlaSerSerGluGluLysCysAlaProLeuAlaThrGluValPheAspGluLysMetGlu 388
Db 122 GCATGTCAGAGGAGAGTGTGCTCTTGGCAACGGAAGTGTTCATGAGAAGATGGAA 181
Qy 389 AlaHisGlnGluValValAlaGluValHisValSerThrValGluLysThrGluGluGlu 408
Db 182 GCCCACCAGAAGTTGTCAGAGGTCCAGCTGAGCACCCTGGAGAGACAGAGGAGGAG 241
Qy 409 GlnGlyGlyGlyGluAlaGluGlyGlyValValValGluGlyThrGlyGluSerLeu 428
Db 242 CAGGAGGAGGAGGAGGCTGAAGGGGGCGTGGTAGAAGGACAGGAGAATCCTTG 301
Qy 429 ProProGluLysLeuAlaGluProGlnGluValProGlnGluAlaGluProAlaGluGlu 448
Db 302 CCCCTCTGAGAAATGCTGCTGAGCCCCAGGAGGTCCCCAGGAAGCTGAGCCTGCTGAGGAG 361
Qy 449 LeuMetLysSerArgGluMetCysValSerGlyLysAspHisThrGlnLeuThrAspLeu 468
Db 362 CTGATGAAGACAGACAGATGTGTCTCTGGAGGAGACACACTCACTGACAGACCTA 421
Qy 469 SerProGluGluLysThrLeuProLysHisProGluGlyIleValSerGluValGluMet 488
Db 422 AGTCTCTGAAGAGAAGCGCTGCCCAAAACCCAGAAAGGCAATTCAGTGGAGTGGAGATG 481
Qy 489 LeuSerSerGlnGluArgIleLysValGlnGlySerProLeuLysLysLeuPheSerSer 508
Db 482 CTGCTCTCTCAGGAAAGAAATCAAGGTACAGGGAAGTCCCTTGAAGAAACTCTTCAGTAGC 541
Qy 509 SerGlyLeuLysLysLeuSerGlyLysLysGlnLysGlyLysArgGlyGlyGlyAsp 528
Db 542 TCAGGCTTAAAGAGAGTGTCTGGGAAGACAGAGGGAACAGAGAGGTGGGGGAGAC 601
Qy 529 GluGluProGlyGluTyrGlnHisIleHisThrGluSerProGluSerAlaAspGluGln 548
|||||
```



```

Db 602 GAAGAGCCGCGAATATCAACACATTCACACGAATCCCGAGAGTCTTAGAGCAG 661
Qy LysGIYgluSerSerAlaSerSerProgluGluprogluGlutThrCysLeuGIuIyS 568
Db 662 AAGGAGAGAGAGCTGCGCGTCCGCCGAGAGCCTGAGAGACCAACGCTGTGGAGAA 721
Qy 569 GIYProLeuGIuAlaProGlnaspGIYgluIaGIuGIuGIYThrThSeraspGIYglu 588
Db 722 GGGCCGCTGGAGACACCCAG-GATGGGAAAGCTGAGGAAGAACTAATTCCGT--GGAGAG 778
Qy LysIYAspXgluGIYIleThrProTrpAlaSerPheLysIYMetValThrProLysIY 608
Db 779 AAGAAGAGGAGG--ATCACTCCCTGGGCATCTTCAAAAAGATGGTGAACCCAGAA 836
Qy ArgValAlaArgProSerGIuSerAspLysGIuGIuGIuGIuGIuGIuGIuGIuGIu 628
Db 837 CCGTC-CGAAGACCTTCTAGAGTGAAGAAGAGAAAGCTGGAGAAAGTCAAGAGCGCC 895
Qy ThrLeuSerSerThrAspSerThrValSerGIuMetGlnaspGIuValIYThrValGIY 648
Db 896 ACCCTGCTCTCCACTGATAGCAAGTGTCAAGAAATGCAAGATGAAGTCAAACTGTGGT 955
Qy 649 GIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIu 668
Db 956 GAGGAACAAAGCCAGAGGAACCAAGCGTAGGGTGAATCTTCAAGTGTGGAGAGCA 1015
Qy 669 LeuIleCysValGIYSerSerIYsIYsIYsIYsIYsIYsIYsIYsIYsIYsIYsIY 688
Db 1016 CTGATTTGTGTGGATCATCCAAAGAAAGAGCAAGGAAGCATCTTCAAGTATTAAGA 1075
Qy 689 GIYGIYProIYThrLeuGIY-GIYAspSerHiIaArgIaGIuGIuAlaSerIYsIYsIY 708
Db 1076 ---GGGCCAAAGACACTGGAGGGGGAGCAGTCAAGAGCAGAGGAGCCAGCAAAACAAA 1132
Qy 708 sGIuIaIaGIYThrAspAlaValProIaSerThrGIuGIuGIuGIuGIuGIuGIuGIu 728
Db 1133 AAGAAGCCGA-ACAGAGCGTGTCTGCGAGACCCAGAGAGCAGACCAAGCGAAAGAA 1191
Qy 728 rSerSerProGIuProAlaGIYSerSerSerGIuGIuGIuGIuGIuGIuGIuGIuGIu 748
Db 1192 TTCCTCACCCGAGCCAGCGGAGAGCCCTTCCAGAGGGAGAGTGTCTCCACTTGGAGTCT 1251
Qy 748 rPheIYsIYsIYsIYsIYsIYsIYsIYsIYsIYsIYsIYsIYsIYsIYsIYsIYsIY 767
Db 1252 ATTTAAAGATTAGTCACTCAAGAAATAATCAAGTCAAACTGAAGAGAGAAAGA-AG 1310
Qy 768 ----GIuAspSerSerVal---GIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIu 785
Db 1311 CCGGAAGGACTTGAAGTGTAGAGAGAGGTTGTCTCACTGAGATCGAACCGGTAGAGAA 1370
Qy 785 uSerTrpValSerIleLysIYsPheIleProGIYArgArgIYsIYsIYsIYsIYsIYsIY 805
Db 1371 ATCTTGCGTTCATTAAAGAAATTCATCCCGGAGCGCGGAAGAAAGGAGATGGGA 1430
Qy 805 s-GIuGIuGIuAlaThrValGIuAspSerGIYProValGIuIleasnGIuAspProIa 825
Db 1431 GGCAAGAACCAAGCACTGTGGAAGACTCAGGGCGCACTGAGATTAATGAGGAGCAAGCTCTG 1490
Qy 825 snValProIaIaValIaProLeuSerGIuThrAspAlaValGIuIaGIuGIuGIuGIu 845
Db 1491 AGTGTCCACCGCGTGTCTGTCTGAGATGATGATCAGATGGAGAGGAGAGATGGAAG 1550
Qy 845 IaGIuGIuIYsanthrGIuLeuProGIuLeuGIuGIuGIuGIuGIuGIuGIuGIuGIu 865
Db 1551 CCGAGGGGAATGGAGAGCTGCCAGCTGTGGGGCTGTGTA---GTGTCCGAGAGACTCA 1607
Qy 865 eIYsIYsIYsIYsIYsIYsIYsIYsIYsIYsIYsIYsIYsIYsIYsIYsIYsIYs 885
Db 1608 GTAAAGCTCTGGTCCACACTGTGATGTGTGAGTCAATGATGGAGCAAGGGCAGTCAACA 1667
Qy 885 eValGIuGIuIYsSerProSerTrpIleSerIaSerValThrGIuProLeuGIuHiIeT 905

```

```

Db 1668 GTGTGGAAGAGCGGTCTCTCTGTGGAATGCCGCTTCCGTAAACAGAACTTGAACACA 1727
Qy 905 hrIaGIYGIuAlaIleMeProProValGIuGIuValIThrGIuLysAspIleIleGIuGI 925
Db 1728 CAGCGGAGAGACCATGCGCACTGTGGAAGAGTCACTGAAAGAAAGACTATTCAGAGAG 1787
Qy 925 IuThrProValIleuThrGIuThrIleuProGIuGIuGIuGIuGIuGIuGIuGIuGIu 945
Db 1788 AAACCTCTGTCTCAACCCAGAGCTTACAGAGGGGTAAAGATGCCATGACAGATGTGTCA 1847
Qy 945 hrSerGIuValAspPheThrSerGIuAlaValIThrAlaThrGIuThrSerGIuAlaLeu 965
Db 1848 CCAGGAAGTGAATTCACCTCAGAACTGTGACAGCCAGAGAACCTCAGAGGCTCTCC 1907
Qy 965 rGIuGIuGIuValIThrGIuAlaSerGIYAlaGIuGIuGIuGIuGIuGIuGIuGIuGIu 985
Db 1908 GTACTGAAGAGATTACCAAGACATCGGGGGCCGAAGAGACCAAGACAGATGTGTCCGAG 1967
Qy 985 aIserGIuLeuThrAspSerProAspThrThrGIuGIuGIuGIuGIuGIuGIuGIuGIu 1005
Db 1968 TTTCCAGACTGACTGACTCCCGAGACACACAGAGAGAAAGCCAGTTGAGAGGTAG 2027
Qy 1005 IuSerGIYValIleuAspThrGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIu 1025
Db 2028 AGGTGTGTCTGATACAGAGAGAGAGAGAGCCCGCAGAGCCAGCTTCCAGAGCCG 2087
Qy 1025 aIaIaAspIYsValIYsGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIu 1045
Db 2088 TTGCAGACAAAGTGAAGAGAGTCCAGGTGCTGCAACCCAGACTGTGCAGAGAAAGG 2147
Qy 1045 IYserIYsIaIleuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIu 1065
Db 2148 GGTCAAAAGCACTGGAGAGAGTGAAGAGGTGAAGAGAGTCCGAAAGTGTGGCTTGG 2207
Qy 1065 IuIYsGIuIYsAspValIleMeProIYsGIYProValGIuGIuAlaGIYAlaGIuIleu 1085
Db 2208 AAGAAGAGAGAGACTTATGCCGAAGAGACCCGTGCAGAGAGCTGTGAGACTTGTG 2267
Qy 1085 IaGIuGIYSerGIuThrGIYGIuAlaThrProGIuSerIleuGIuValProGIuValIThr 1105
Db 2268 CACAGGCTGTGAGACTGAGACAGGCTACTCCAGAGAGCTTGAAGTCTCGAAGTCAAG 2327
Qy 1105 IaAspValAspHiIaValaIaThrCysGIuGIuValIleYsLeuGIuGIuIleuMeTGIuGIu 1125
Db 2328 CAGATGTAGACCATGTGCCAGTGCAGGTATCAAGCTTCAGACGTGATGGAACAGG 2387
Qy 1125 IaValaIaProGIuSerSerGIuThrIleuThrAspSerGIuThrAsnGIYSerThrPro 1145
Db 2388 CCGTGGCCCTGAGTCACTCCGAACCTTGAACAGACAGTGAACAAATGGAAGCACTCCCT 2447
Qy 1145 euIaIaAspSerAspThrIaIaAspGIYThrGIuGIuGIuGIuGIuGIuGIuGIuGIuGIu 1165
Db 2448 TAGCAGATTCAAGACACTGCAATGAGGAGACAGCAAGAAAGAAACCAATGACAGCCAGACA 2507
Qy 1165 eIYsIaIaThrAlaIaValaIaArgIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIu 1185
Db 2508 GTAAAGCCACTGCACGTGTGAGGCGTCAAGGTCAAGAAAGAGAGCGGCTACTGTCTC 2567
Qy 1185 IuIYsGIuGIuProSerThrIleuProAsnAsnValProAlaGIuGIuGIuGIuGIuGIu 1205
Db 2568 AGAAAGAGAGGCTTCCAGACATCACTAATATGTTCCAGCCAGAGAAACATGGGGAAG 2627
Qy 1205 IuProGIYArgAspValIleuGIuProThrGIuGIuGIuGIuGIuGIuGIuGIuGIuGIu 1225
Db 2628 AACAGAGAAAGATGTCTTGAACCTTACAGACAGCAAGAGACTGTGTCTGAGCGGCCG 2687
Qy 1225 aIleuAlaIYsThrGIuValaIaGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIu 1245
Db 2688 TCTGGCAAAAGACTGAGGTGTGCTCAAGAGGGTGAAGTGTGACTGTGTGATGGAAGAAAG 2747
Qy 1245 aIYsGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIu 1265
Db 2748 TCAGAGAGAGAGAGAGGTGTTGTACCTTGAACCCAGACAGTCAAAAGGCTGTGATG 2807

```


Oy 39 AlaGlyAspProAlaAspAla-----AspProAlaThrLysLeuProGluLysAsnGly 56
 Db 312 ACCGGGAGCCCGCATGCTGCTCGGACCCCGCCACCAAGCTCTCAGAAAGATAGT 371
 Oy 57 GlnLeuSerSerValAsnGlyValAlaGluGlnGlyAspValHisValGlnGluLysn 76
 Db 372 CAGCTGTCACCATCAATAGGCTAGCTAGCAAGATGAGCTCAGCTCCAGGAGGTGAC 431
 Oy 77 GlnGluGlyGln-----GlnGlu 82
 Db 432 CTAATATGCGCAGAAAGAGACCCCTGAACGCTCAGAGACCCCTAAACAGCCAGGAGAGAA 491
 Oy 83 GluValValAspGluAspValGlyGlnArgGluSerGluAspValArgGluLysAspArg 102
 Db 492 GAAGTCATTGTACACAGAGTTTGACAGAGAGACTCTGAAGATGTGAGCGAAAGAGACTCC 551
 Oy 103 ValGluGluMetAlaAlaAsnSerThrAlaValAlaGluAspLleThrLysAspGlyGlnGlu 122
 Db 552 GATAAAGAGATGGCTACTAAGTCAAGCGTTGTTCAGACATCAGATGATGGGAGAG 611
 Oy 123 GluThrSerGluLleIleGluGlnLleProAlaSerGluAsnAsnValGluGluMetVal 142
 Db 612 GAGACACCCGAAATATATGACACAGATTCCTTCCAGAAAGCAATTTAGAGAGCTTAACA 671
 Oy 143 GlnProAlaGluSerGlnAlaAsnAspValGlyPheLysLysValPheLysPheValGly 162
 Db 672 CAACCCATCGAGTCCAGGCTAATGATATTTGATTTAAGAGGTGTTTAAAGTTGTGGC 731
 Oy 163 PheLysPheThrValLleLysLysAspLysAsnGluLysSerAspThrValGlnLeuLeuThr 182
 Db 732 TTTAAATTCACGTGAAAAAGATTAAGACAGAGAAAGCTGACACGTGTCCAGCTACTACT 791
 Oy 183 ValLysLysAspGluGlyGluGluLysLysValGluLysValGlyValAlaGlyAspHisGlnGlu 202
 Db 792 GTGAAGAAAGATGAAGGGGAGGAGCA-----GCGGGGGCTGGGCAACACAGAGAC 842
 Oy 203 ProSerValGluThrAlaValAlaGlyLysLysLysLysLysLysLysLysLysLysLys 222
 Db 843 CCCACCTT-----GGGGCTGGAAGAGCAGCATCCAAAGAAAGCAACCAATCT 896
 Oy 223 ThrGluLysGlnGluGlyThrLeuLysGlnGluGlnGlnGlnGlnGlnGlnGlnGln 242
 Db 897 ACAGAGAAACCCGAGAGAGCCTTGAAGCGTGAAGCAAGCAAGCAAGCAAGCAAGCA 956
 Oy 243 AlaGluSerAspGlnAlaAlaGluGluGluLysLysLysLysLysLysLysLysLysLys 262
 Db 957 GCCGATCTGGCCAGAGCAAGT--GAGGATGCAAGAGAGAGAGAGAGAGAGAGAGAA 1013
 Oy 263 LysGluProThrLysSerProGluSerProSerProValAsnSerGluThrThrSer 282
 Db 1014 AAAGAACCTAGCAAGCTGTCGACATCTCCGACTAGTCCCGTGAACAGTGAACAGATCA 1073
 Oy 283 SerPheLysLysPhePheThrHisGlyTTPAlaGlyTTPArgLysLysThrSerPheLys 302
 Db 1074 ACCCTCAAAAAATTTCTTCACTCAAGGTTGGCGCGCTGCGCAAAAAAGCCAGTTTCAG 1133
 Oy 303 LysSerLysGluAspAspLeuThrAlaGluLysArgLysGluGlnGluLysLys 322
 Db 1134 AAGCCGAAAGAGATGATGAGAGCTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAA 1193
 Oy 323 ValAspGluGluGluLysGluLysThrGluProAlaSerGluLys----- 337
 Db 1194 GTAGGACACGAG 1253
 Oy 338 -----GlnGluProAlaGluAspThrAspGlnAlaArgLeuSerAlaAspLys 353
 Db 1254 CAAGCCCAACCCACAGAGAGCGGCAAGAAAGTCCCAAGAGCCCGGTTATCACCTGAATAT 1313
 Oy 354 GluLysValGluLeuProLeuGluAspGlnValGlyAspLeuGluLysSerGluGlu 373
 Db 1314 GAGAAAGTTGAGCTGCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1373

Oy 374 LysCysAlaProLeuAlaThrGluValPheAspGluLysMetGluAlaHisGln---Glu 392
 Db 1374 AAACCTGCTCCGTGGGAGACAGAAAGTGTTCATGAGAAATAGAAATGCCACCAAGAGAG 1433
 Oy 393 ValValAlaGluValHisValSerThrValGluLysThrGluGluGluGluGlyGly 412
 Db 1434 GTTGTGCCCAAGATCCACGTACGACCCGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1484
 Oy 413 GlyLysAlaGluGlyGlyValValValGluGlyThrGlyLysSerLeuProGluLys 432
 Db 1485 -----AAACCGAGGTGAGAGAAACAGAGAGGTCTGTGTCCAGCTGAAGAA 1529
 Oy 433 LeuAlaGluProGlnGluValAlaProGlnGluAlaGluProAlaGluGluLeuMetLysSer 452
 Db 1530 TTGTTGAATATGATGACAGAACTTCAGAGAACCCAGACCTTCAGAGAGCTGTAAAGCTC 1589
 Oy 453 ArgGluMetCysValSerGlyGlyAspHisThrGlnLeuThrAspLeuSerProGluGlu 472
 Db 1590 AAGAGACGTGTGTTCGAGAGAGACCTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1649
 Oy 473 LysThrLeuProLysHisProGluGlyLleValSerGluValGluMetLeuSerGln 492
 Db 1650 AAGTGTCTGCCAAACCCCGGAGAGGCTGTGAGTGAAGTGAAGATGCTGTATCAAG 1709
 Oy 493 GluArgLleLysValGlnGlySerProLeuLysLysLeuPheSerSerGlyLysLys 512
 Db 1710 GAGAGAAATGAAGGTGCGAGGAGATCCACTAAAGAAAGCTTTTACAGACATCGGCTTAAA 1769
 Oy 513 LysLeuSerGlyLysLysGluLysGlyLysArgGlyGlyGlyGlyGlyGluGluProGly 532
 Db 1770 AAGCTTTTGAGAAAGAAACAGAAAGGAGAAAGAA--GAGAGAGAGACAGAGAAATCAAGG 1826
 Oy 533 GluTyrglnHisIleHisThrGluSerProGluSerAlaAspGluGlnLysGlyLys 552
 Db 1827 GAGACACACTAGTTCACAGCGGATTCCTCGGACAGCCAGAGAGAGAGAGAGAGAGAG 1886
 Oy 553 SerAlaSerSerProGluGluProGluGluThrThrCysLeuGluLysGlyProLeuGlu 572
 Db 1887 TCTGCTCATCTCCCTGAGAGACCCGAGAGATCAAGTCTGTGAGAAAGGGCTTAGCCAG 1946
 Oy 573 AlaProGluAspGlyGluAlaGluGluGlyThrThrSerAspGlyGluLysArgGlu 592
 Db 1947 GTGACAGCGAGATGGGAGAGCTGAAGAGAGAGACTTCTCGATGAGAGAGAGAGAGAA 2006
 Oy 593 GlyLleThrProTTPAlaSerPheLysLysMetValThrProLysArgValArgArg 612
 Db 2007 GGTGTCACTCCCTGGGATTCATTCAAAAGATGTGAGAGCCCAAGAGAGCGTTAGAGCG 2066
 Oy 613 ProSerGluSerAspLysGluGluGluLysValLysSerAlaThrLeuSerSer 632
 Db 2067 CCTTCGAAAGTGAATAAGAGATGAGCTGAGCAAGGTCAAGAGCGCTACTTGTCTTCC 2126
 Oy 633 ThrAspSerThrValSerGluMetGlnAspGluValLysThrValGlyGluGluLys 652
 Db 2127 ACCGAGACACAGCCTCTGAATGCAAGAGAAATGAAGAGAGAGAGAGAGAGAGAGAG 2186
 Oy 653 ProGluGluProLysArgArgValAspThrSerValSerTTPGluAlaLeuLecysVal 672
 Db 2187 CCGAGAGAACCAAGCCGAGAGGTGATCTTCAGATCTTGGAAAGCTTAAATTTGTGTG 2246
 Oy 673 GlySerSerLysLysArgAlaArgLysAlaSerSerSerAspAspGluGlyProArg 692
 Db 2247 GGATCATCCAGAAAG 2306
 Oy 693 ThrLeuGlyGlyAspSerHisArgAlaGluGluLysSerLysAspLysGluAlaGlyThr 712
 Db 2307 GCATATGAG 2366
 Oy 713 AspAlaValProAlaSerThrGlnGluGlnAspGlnAlaGlnGlyLysSerSerProGlu 732
 Db 2367 GACGGGATCTTGTGCTGTCTCCAGAGACATGATCCAGGGCAGAGAGATCTTCCCCGAG 2426
 Oy 733 ProAlaGlySerProSerGluGlyGlyValSerThrTTPGluSerPheLysArgLeu 752

Db 2427 CAAGCTGGAAGCCCTCCGAAGGGAGGGCGTTTCCACCTGGGAGTCATTATAAAGGTTA 2486
Qy 753 ValThrProArgLysSerLysSerLysLeuGluGluLysAlaGluAAsp-----ValaLaThrCys 1113
Db 2487 GTACAGCCCAAGAAAAATCAAGCTCAAGCTGGAAGAGAAAGCGAAGACTCCATAGCT 2546
Qy 770 ---SerSerValGluGlnLeuSerThrGluIleGluProSerArgGluGluSerTrpVal 788
Db 2547 GGGTCTGGTGTAGAACATTCACCTCCAGACACTGAACCCGGTAAAGAAATCTCGGGTC 2606
Qy 789 SerIleLysLysPheIleProGlyArgGlyLysArgAlaAspGlyLysGlnGluGln 808
Db 2607 TCAATCAAGAAAGTTTATCTCGACGAAGGAAGAAAGGCCAGATGGGAACAAGAACAA 2666
Qy 809 AlaThrValGluAspSerGlyProValGluIleAsnGluAspAspProAsnValProAla 828
Db 2667 GCCCCTGTGAAGACGCGAGGCCCAACAGGGGCCAACGAAGATGACTCTGTATGTCCTCGGCC 2726
Qy 829 ValValProLeuSerGluTyrAsnAlaValGluArgGluLysMetGlu-----Ala 845
Db 2727 GTGGTCCCTCTGTCTGAGTATGATGCTGTGAAAGGGAGAAAATGAGGCACAGCAAGCC 2786
Qy 846 GlnGlyAsnThrGluLeuProGlnLeuGlyAlaValTyrValSerGluGluLeuSer 865
Db 2787 CAAAAAGCGCAGAGCAGCCGACAGAGCCACACTGAGGTGTCCAGAGACTCAGC 2846
Qy 866 LysThrLeuValHisThrValSerValAlaValIleAspGlyThrArgAlaValThrSer 885
Db 2847 GAGAGTCAGGTTCATATGATGCGCAGCAGCTGCTGACGGGACGAGGCGACTACCAT 2906
Qy 886 ValGluGluArgSerProSerTrpIleSerAlaSerValThrGluProLeuGluHisThr 905
Db 2907 ATTGAAGAAAGGTCTCCTCTCTGTGATATCTGCTTCAGTGACAGAACCTCTTGAACAAGTA 2966
Qy 906 AlaGlyGluAlaMetProProValGluGluValThrGluLysAspIleIleAla---Glu 924
Db 2967 GAAGCTGAAGCCGACCTGTTAACTGAGGAGGTATGTGAAGAGAGAAATGTGCAGAGAA 3026
Qy 925 GluThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHisAspAspMetVal 944
Db 3027 GAACCCCCACCGTTACTGAACCTCTGCCAGAGAACAGAGAGGCCCGGGCGACACGGTC 3086
Qy 945 ThrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeu 964
Db 3087 GTTAGTGAGGCGGAATTGACCCCGAAGCATCTGCTGTAAGAGACCACAGAAATGGTGTACGA 3146
Qy 965 ArgThrGluGluValThrGluAlaSerGlyAlaGluThrThrAspMetValSerAla 984
Db 3147 GGTTCGNAAGAGAACCGAAGCATCTGCTGCTGAAGAGACCACAGAAATGGTGTACGA 3206
Qy 985 ValSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGluVal 1004
Db 3207 GTCTCCAGTTAACCGACTCCACAGACACACAGAGAGGCCACTCCGGTGCAGAGGTG 3266
Qy 1005 GluSerGlyValLeuAspThrGluGluGluGluArgGlnThrGlnAlaIleLeuGlnAla 1024
Db 3267 GAAGGTGGCGTACCTGACATAGAGAGACAGAGAGGGCGGACTCAAGAGGTCTCTCCAGCCA 3326
Qy 1025 ValAlaAspLysValLysGluGluSerGlnValProAlaThr----- 1038
Db 3327 GTGCAGAAAAGTGAAGAGGAATCCAGCTGCTCGCACCCGGTGGGCCAGAGATGTG 3386
Qy 1039 ---GlnThrValGlnArgThrGlySerLysAlaLeuGluLysValGluGluValGluGlu 1057
Db 3387 CTTTCAGCTGTGCAGAGACA-----GAGGCAGAAAAGACCAGAAAGAG 3428
Qy 1058 AspSerGluValLeuAlaSerGluLysGluLysAspValMetProLysGlyProValGln 1077
Db 3429 CAGGCTGAAGCGTCGGGTCTGAAGAAAGAGACGGATGTAGTGTGAAAGTAGATGCTCAG 3488
Qy 1078 GluAlaGlyAlaGluHisLeuAlaGlnGlySerGluThrGlyGlnAlaThrProGluSer 1097

Db 3489 GAGGCAAAAACTGAGCCTTTTACACAAGGGAAGGTGGTGGGGCAGACCACCCCGAGAAAGC 3548
Qy 1098 LeuGlu---ValProGluValThrAlaAspValAspHis-----ValaLaThrCys 1113
Db 3549 TTTGAAAAGAGCTCCTCAAGTCACAGAGAGCATAGAGTCCAGTCCAGCTTGTAAACCACTGT 3608
Qy 1114 GlnVal-----IleLysLeuGlnGlnLeu---MetGluGlnAlaValAla 1127
Db 3609 CAAGCCGAAACCTTAGCTGGGTAAATAATCAAGAGATGGTGTATGGAACAGGCTATCCCC 3668
Qy 1128 ProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrProLeuAlaAsp 1147
Db 3669 CTTGACTCGGTGGAACCCCTCAGACAGTGAAGTGTGAAGACCCCGCTGATGAGCCGAC 3728
Qy 1148 SerAspThrAlaAspGlyThrGlnGlnAspGluThrIleAspSerGlnAspSerLysAla 1167
Db 3729 TTTGAGCACCAGGCACAACCCAGAAAGACGAGATTGTGGAATCCATGAGGAGAATGAG 3788
Qy 1168 ThrAlaAlaValArgGlnSerGlnValThrGluGluGluAlaAlaThrAlaGlnLysGlu 1187
Db 3789 GTCCGATCTGGTATCCAGTCCAGGGGGCACAGAAGCAGGAGCAGTTCCTGCACAGAAAGAG 3848
Qy 1188 GluProSerThrLeuProAsnAsnValProAlaGlnGluGluHisGlyGluGluProGly 1207
Db 3849 AGGCCTCCAGCA---CCTTCAGTCTTTGTTCCAGGAAGAACTAAAGAACATCAAG 3905
Qy 1208 ---ArgAspValLeuGluProThrGlnGlnLeuThrAlaAlaValProValLeu 1226
Db 3906 ATGGAAGACACTCTAGAGCATACAGATAAAGAGGTCTCAGTGGAAACTGTATCCATTCTG 3965
Qy 1227 AlaLysThrGluValGlyGlnGluGlyGluValAspTrpLeuAspGlyGluLysValLys 1246
Db 3966 TCAAAGACTGAGGGGACTCAA-----GAGGCTCACCAGTATGCTGATGAGAAACCAAA 4019
Qy 1247 -----GluGluGlnGluValPheValHisSerGly-----ProAsnSer 1259
Db 4020 GAGCTACCATTTTTCGAAGGACTTGAGGGGTCTATAGACACAGGCATTAACAGTCAGTCGG 4079
Qy 1260 GlnLysAlaAlaAspValThrTyAspSerGluValMetGlyValAlaGlyCysGlnGlu 1279
Db 4080 GAAAGGTCACTCAAGTTGCCCTTAAAGGTGAAGGGACAGAGAGCTGAATGTAAAAAG 4139
Qy 1280 LysGluSerThrGluValGlnSer-----LeuSerLeuGluGluGlyGlu 1294
Db 4140 GATGATGCTCTTGAACCTGCAGAGTCAAGTCAAGTCTCTCCATCCCTCCGCTGGAGAGAG 4199
Qy 1295 MetGluThrAspValGluLysGluLysArgGluThrLysProGluGlnValSerGluGlu 1314
Db 4200 ATGGTAGTTTCAAGTCAAGGGGAGAAAACAGAACGAGCCAAACCCATGTGAATGAAGAG 4259
Qy 1315 Gly---GluGlnGluThrAlaAlaProGluHisGluGlyThrTyGlyLysProValLeu 1333
Db 4260 AAGCTTGAGCACAACAGCTGTTACCGTATCTGAAGAGGTCTAGTAAGCAGCTCCTCCAG 4319
Qy 1334 ThrLeuAspMetProSerSerGluArgGlyLysAlaLeuGlySerLeuGlyGlySer--- 1352
Db 4320 ACAGTGAATGTGCCCATCATAGATGGGCAAGAGAAAGTCAAGCTTGGAAAGAAAGCCCT 4379
Qy 1353 ProSerLeuProAspGlnAspLysAlaGlyCysIleGluValGlnValGlnSerLeuAsp 1372
Db 4380 CCTCCCTGCTAGGTCAAGAGGAGGAGGTATGCACCAAAATTCAGATTCAGAGCTCTGAG 4439
Qy 1373 ThrThrValThrGlnThrAlaGluAlaValGluLysValIle-----GluThrValVal 1390
Db 4440 GCATCATCTCTACAGCGGCTGCAGAGGAGAAAAGGTCTTAGAGGAAATGCCAAC 4499
Qy 1391 IleSerGluThrGlyGluSerProGluCysValGlyAlaHisLeuLeuProAlaGluLys 1410
Db 4500 ATTTAGAAACAGGTGAAAGCTTGGAGCCTGCAGGTGCACATTTAGTTCTTGGGAAGAAA 4559
Qy 1411 SerSerAlaThrGlyGlyHisTrpThrLeuGlnHisAlaGluAspThrValProLeuGly 1430
Db 4560 TCCTCTGAAAAAATGAAGACTTTTGGCGCTCATCCAGGGGGAAGATGCTGTGCCACAGGG 4619

Db 1251 CAAGCCACCCACAGGAGCGGCAGAAAGTCCCCACGAGCCCGGTTATCAGCTGAATAT 1310
Qy 354 GluIysValGluLeuProLeuGluAspGlnValIcIyAspLeuGluAlaSerSerGluGlu 373
Db 1311 GAGAAAGTTGAGCTGCCCTCAGAGGAGCAAGTCTAGTGGCTCGCAGGGACCTTCTGAAGAG 1370
Qy 374 LysCysAlaProLeuAlaThrGluValPheAspGluLysMetGluAlaHisGln---Glu 392
Db 1371 AAACCTGCTCGTGGCGACAGAAAGTGTGTGATGAGAAATATAGAGTCCACCAGAAAGAG 1430
Qy 393 ValValAlaGluValHisValSerThrValGluLysThrGluGluGluGlnGlyGly 412
Db 1431 GTTGTGCCAGGATCCACGTCAGCACCGTGGAGGAGAGAACCGAAGAGCAG----- 1481
Qy 413 GlyGluAlaGluGlyGlyValValGluGlyThrGlyGluSerLeuProGluLys 432
Db 1482 -----AAAACGGAGGTGGAAGAACACAGCAGGGTCTGTGCCAGCTGAAGAA 1526
Qy 433 LeuAlaGluProGlnGluValProGlnGluAlaGluProAlaGluGluLeuMetLysSer 452
Db 1527 TTGTTGGAATGGATGCAGAACCTCAGAAAGCCGAACCTGCCAAGGAGCTGGTGAAGCTC 1586
Qy 453 ArgGluMetCysValSerGlyGlyAspHisThrGlnLeuThrAspLeuSerProGluGlu 472
Db 1587 AAAGAAACGTTGTTTCCGAGGAGGACCTACACAGGAGCTGACCTCAGTCTGTATGAG 1646
Qy 473 LysThrLeuProLysHisProGluGlyIleValSerGluValGluMetLeuSerSerGln 492
Db 1647 AAGGTGCTGTCCAAACCCCGAAGGCGTGTGAGTGAGGTGGAATGCTGTATCATCAG 1706
Qy 493 GluArgIleLysValGlnGlySerProLeuLysLysLeuPheSerSerGlyLeuLys 512
Db 1707 GAGAAATGAAGTGCAGGGAAGTCCACTAAAGAGCTTTTACCAGCACTGGCTTAAAA 1766
Qy 513 LysLeuSerGlyLysLysGlnLysGlyLysArgGlyGlyGlyGlyAspGluGluProGly 532
Db 1767 AAGCTTCTGGAAAGAACACAGAAAGGAAAGA---GGAGGAGAGACGAGGATCAGG 1823
Qy 533 GluThrGlnHisIleHisThrGluSerProGluSerAlaAspGluGlnLysGlyGluSer 552
Db 1824 GAGCAGACTCAGGTTCACGCGGATTTCTCCGACAGCCAGGAGGAGCAAAAGGCGAGAGC 1883
Qy 553 SerAlaSerProGluGluProGluGluThrThrCysLeuGluLysGlyProLeuGlu 572
Db 1884 TCGCTCATCCCTGAGGAGCCGAGGAGATCAGCTGTCTGAAAGAGGCTTAGCCGAG 1943
Qy 573 AlaProGlnAspGlyGluAlaGluGluGlyThrThrSerAspGlyGluLysLysArgGlu 592
Db 1944 GTCAGCAGGATGGGAAGCTGAAGAGGAGCTACTTCCGATGGAGAGAAAGAGAA 2003
Qy 593 GlyIleThrProTrpAlaSerPheLysLysMetValThrProLysLysArgValArg 612
Db 2004 GGTGTCACTCCCTGGGCATCATTCAAAGATGTGTGACGCCCAAGAGCGTGTAGACGG 2063
Qy 613 ProSerGluSerAspLysGluGluGluLeuGluLysValLysSerAlaThrLeuSerSer 632
Db 2064 CCTTCGAAAGTATAAAGAGATGAGCTGGACAAGGTCAAGAGCGCTACCTTGTCTTCC 2123
Qy 633 ThrAspSerThrValSerGluMetGlnAspGluValLysThrValGlyGluGluGlnLys 652
Db 2124 ACCGAGAGCAGCCCTCTGAATGCAAGAGAAATGAAGGGAGCGTGGAGAGCCAAAG 2183
Qy 653 ProGluGluProLysArgValAspThrSerValSerTrpGluAlaLeuIleCysVal 672
Db 2184 CCGAAGAACCAAGCCAAAGTGGATACCTCAGTATCTTTGGGAAGCTTTAATTTGTGTG 2243
Qy 673 GlySerSerLysLysArgAlaArgLysAlaSerSerSerAspGluGlyGlyProArg 692
Db 2244 GGAATCATCCAAAGAAAGAGCAAGGAGGAGGTCTCTTCTGATGAGGAAGGGGACCAAAA 2303
Qy 693 ThrLeuGlyGlyAspSerHisArgAlaGluGluAlaSerLysAspLysGluAlaGlyThr 712
Db 2304 GCATGGGAGGAGACCCACAGAAAGCTGTATGAGCCCGGAAAGACAAAGAGACGGGACA 2363

Qy 713 AspAlaValProAlaSerThrGlnGluGlnAspGlnAlaGlnGlySerSerSerProGlu 732
Db 2364 GACGGGATCCTGCTGTTCCCAAGAACATGATCCAGGGCAGGGAAGTCTCTCCCGGAG 2423
Qy 733 ProAlaGlySerProSerGluGlyGluGlyValSerThrTrpGluSerPheLysArgLeu 752
Db 2424 CAAGCTGGAAAGCCCTACCGAAGGGAGGCGTTCACCTGGGAGTCATTTAAAGGTTA 2483
Qy 753 ValThrProArgLysLysSerLysLeuGluGluLysAlaGluAsp----- 769
Db 2484 GTCAGCCCAAGAAAAAATCAAGTCCAAGTGGAAAGAGAAAAAGCAGACTCCATAGCT 2543
Qy 770 ---SerSerValGluGlnLeuSerThrGluIleGluProSerArgGluGluSerTrpVal 788
Db 2544 GGTCTCGTGGTGTAGAACATTCCTCCAGACACTGAACCCCGTAAAGAAAGAAATCCTGGT 2603
Qy 789 SerIleLysLysPheIleProGlyArgArgLysLysLysAlaAspGlyLysGlnGluGln 808
Db 2604 TCAATCAAGAAAGTTTATCTCTGGACGAAGAGAAAGGCCAGATGGGAAACAAAGAACAA 2663
Qy 809 AlaThrValGluAspSerGlyProValGluIleAsnGluAspAspProAsnValProAla 828
Db 2664 GCCCTGTTGAAGACCCAGGGCCCAACAGGGGCCAACGAAAGATGACTCTGATGTCCCGGC 2723
Qy 829 ValValProLeuSerGluTyrAsnAlaValGluArgGluLysMetGlu-----Ala 845
Db 2724 GTGGTCCCTCTGTCTGATGATGCTGTAGAAAGGAGAGAAATGGAGGCACAGCAAGCC 2783
Qy 846 GlnGlyAsnThrClnLeuProGlnLeuGluAlaValTyValSerGluGluLeuSer 865
Db 2784 CAAAAGGCGCAGAGCAGCCCGCAGAGAGGAGGAGGCTGAGGTGTCCAAAGGAGCTCAGC 2843
Qy 866 LysThrLeuValHisThrValSerValAlaValIleAspGlyThrArgAlaValThrSer 885
Db 2844 GAGAGTCAGTTTATGATGAGCAGCAGCTGTCTGACGGGACGAGGAGGAGTACCAT 2903
Qy 886 ValGluGluArgSerProSerTrpIleSerAlaSerValThrGluProLeuGluHisThr 905
Db 2904 ATTGAAGAAAGGTCTCTCTTGGATATCTGCTTCACTGACAGAACCTCTTGAAACAAGTA 2963
Qy 906 AlaGlyGluAlaMetProProValGluGluValThrGluLysAspIleAla---Glu 924
Db 2964 GAAGCTGAAGCGCCACTGTTAACTGAGGAGGTATTGGAAGAGAGAGTAATTCAGAAGAA 3023
Qy 925 GluThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHisAspAspMetVal 944
Db 3024 GAACCCCCACGGTACTGAACTCTGCCAGAGAACAGAGAGCCCGGGGCGACACGGTC 3083
Qy 945 ThrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeu 964
Db 3084 GTTAGTGAGCGGAATTGACCCCGAAGCTGTGACAGCTGCAGAAACTGCAGGGCCATTG 3143
Qy 965 ArgThrGluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSerAla 984
Db 3144 GGTTCGAAAGAGGAACCGAAGCATCTGCTGCTGAAGAGACCCACAGAAATGTTGTGCTCA 3203
Qy 985 ValSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGluVal 1004
Db 3204 GTCTCCAGTTTAAACCGACTTCCCGACACCCAGAGAGGCCACTCCGGTGCAGGAGGTG 3263
Qy 1005 GluSerGlyValLeuAspThrGluGluGluArgGlnThrGlnAlaIleLeuGlnAla 1024
Db 3264 GAAGGTGGCGTACCTGACATAGAGAGCAAGAGAGGCGGACTCAAGAGGTCTCTCCAGGCA 3323
Qy 1025 ValAlaAspLysValLysGluGluSerGlnValProAlaThr----- 1038
Db 3324 GTGGCAGAAAAAGTGAAGAGGAATCCAGCTCCCTGGCACCCGCTGGCCAGAGAAGATGTG 3383
Qy 1039 ---GlnThrValGlnArgThrGlySerLysAlaLeuGluLysValGluGluValGluGlu 1057
Db 3384 CTTTACGCTGTGCAGAGACA-----GAGCAGAAAGAACCCAGAAAGAG 3425

Db 2244 GCATCATCCAGAAAAAGCAAGAGAGAGAGCTCTCTTCATGATGAGGAAGGGGACCCAAA 2303
 QY 693 ThrLeuGlyIAspSerHisArgAlaGluIValAspLeuAspIleGluAlaGlyThr 712
 Db 2304 GCATATGGGAGAGACACACAGAAAGCTGATGAGCCGGAAGAAAGACAAAGAGACGGGGACA 2363
 QY 713 AspAlaValProAlaSerThrGlnGluGlnAspGlnAlaGlnGlySerSerSerProGlu 732
 Db 2364 GACGGGATCTCTTGTCTGACCAAGAACATGATCCAGGGCAGGGAAGTTCTCCCGGAG 2423
 QY 723 ProAlaGlySerProSerSerGluGluGluIValSerThrTrpGluSerPheIleArgLeu 752
 Db 2424 CAACTGGGAAGCCCTTACCAAGAGGAGGGCGCTTTCACCTGGGAGCTTTAAAGGTTA 2483
 QY 753 ValThrProArgIleLysSerIleLysSerIleLysGluIleLysAlaGluAsp 769
 Db 2484 GTACAGCCCAAGAAAAATCAAGATCCAAAGTCCAAAGAGAAAAAGCGAAGATCTCCATGCT 2543
 QY 770 ---SerSerValGluGluLeuSerThrGluIleGluProSerArgGluGluSerThrVal 788
 Db 2544 GGGCTGGTGTGAAACATTCACCTCCAGACACTGAACCCGGTAAAGAAATCTGGGTC 2603
 QY 789 SerIleLysPheIleProGlyArgGlyLysArgAlaAspGlyLysGluGln 808
 Db 2604 TCAATCAAGAGATTATCTCTGACCAAGAGAAAGAACGCCAGATGGGAAACAAAGACA 2663
 QY 809 AlaThrValGluAspSerGlyProValGluIleAsnGluAspAspProAsnValProAla 828
 Db 2664 GCCCCTGTTGAAGACGACAGGGCCAAAGGGGGCCAAAGAAATGATCATCTGATGCTCCGGCC 2723
 QY 829 ValValProLeuSerGluIleValAsnAlaValGluArgGluLysMetLeu 845
 Db 2724 GTGGTCCCTCTGTGATGATGATGCTGTGAAAGGGAAGAAATGAGGACACCAAGGCC 2783
 QY 846 GlnGlyAsnThrGluLeuProGlnLeuLeuGlyAlaValTyraIleSerGluGluLeuSer 865
 Db 2784 CAAAAAGGCGCAGACGCCCGACGAAAGGCCCACTGAGGTGTCCAAAGAGCTCAGC 2843
 QY 866 LysThrLeuValHisThrValSerValAlaValIleAspGlyThrArgAlaValThrSer 885
 Db 2844 GAGAGTCAGATTATGATGATGAGCAGCTGTGCTGACGAGGACGAGGACAGTACATT 2903
 QY 886 ValGluGluArgSerProSerThrIleSerAlaSerValThrGluProLeuGluIleThr 905
 Db 2904 ATTGAAGAAAGGCTCTCTCTTGGATATCTGCTTCACTGACAAACCTCTTGAACAAGTA 2963
 QY 906 AlaGlyGluAlaMetProProValGluGluValThrGluLysAspIleIleAla---Glu 924
 Db 2964 GAACTGAAGCCCGCACTGTTAACTGAGAGGTTATGAAAGAGAAAGTAATTGCAAGAGA 3023
 QY 925 GlnThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHisAspAspMetVal 944
 Db 3024 GAACCCCGCAGGTTACTGAACCTCTGCCAGAAACAGAGAGGCCCGGGGGCGACAGCTC 3083
 QY 945 ThrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeu 964
 Db 3084 GTTAGAGGAGCGGAATGTGACCCCGAAGCTGTGACAGCTGCAGAAACCTGAGGCCCATTT 3143
 QY 965 ArgThrGluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSerAla 984
 Db 3144 GGTTCGGAAGAGAACCGCAAGCATCTGCTGTAAGAGACCAAGAAATGGTGTACAGCA 3203
 QY 985 ValSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGluGluVal 1004
 Db 3204 GTTCCCAAGTTAACCACTCCCAAGACACACAGAGAGGCCCATCTCGGTGACGAGGGTG 3263
 QY 1005 GluSerGlyValLeuAspThrGluGluGluGluArgGlnThrGlnAlaIleLeuGlnAla 1024
 Db 3264 GAAGGTGGGCTACCTGACATAGAGAGCAAGAGAGCGGACTCAAGAGGTCTCCAGGCA 3323
 QY 1025 ValAlaAspLysValLysGluGluSerGlnValProAlaThr----- 1038
 Db 3324 GTGGCAGAAAAGTGAAAGGAAATCCAGCTGCTGGCACCGGTGGGCCAGAAAGATGTG 3383

QY 1039 ---GlnThrValGlnArgThrGlySerLysAlaLeuGluLysValGluGluValGluGlu 1057
 Db 3384 CTTCAGCCTGTGCAGAGAGCA-----GAGGCAAGAAAGACCAAGAAAG 3425
 QY 1058 AspSerGluValLeuAlaSerGluLysGluLysAspValMetProLysGlyProValGln 1077
 Db 3426 CAGGTGAAGGCTGGGCTTAAGAAAGACGAGATGATGTTGAAAGTGAATGATGCTCAG 3485
 QY 1078 GluAlaGlyAlaGluHisLeuAlaGlnGlySerGluThrGlyGlnAlaThrProGluSer 1097
 Db 3486 GAGGCAAAAACAGACCTTTTACACAAGGGAAGGTGGGGGACAGACACCCCGAAGAAC 3545
 QY 1098 LeuGlu---ValProGluValThrAlaAspValAspHis-----ValAlaThrCys 1113
 Db 3546 TTGGAAGAAAGCTCTCAAGTACAGAGACATTAAGTCCAGTGAAGCTTGTAAACCATTTGT 3605
 QY 1114 GlnVal-----IleLysLeuGlnGlnLeu---MetGluAlaValAla 1127
 Db 3606 CAAAGCCGAAACCTTACCTGAGGCTTAATCACAGAGATGATGAGACAGGCTATCCCC 3665
 QY 1128 ProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrProLeuAlaAsp 1147
 Db 3666 CCTGACTCGGTGGAAACCCCTACAGACAGTGAAGTGAAGACCCCGGTAGCCGAC 3725
 QY 1148 SerAspThrAlaAspGlyThrGlnGlnAspGluThrIleAspSerGlnAspSerLysAla 1167
 Db 3726 TTTCAGCAGCACAGGACCAACCCAGAAAGACAGATTTGGAAATCCATAGAGAGATGAG 3785
 QY 1168 ThrAlaAlaValArgGlnSerGlnValThrGluGluGluAlaIleThrAlaGlnLysGlu 1187
 Db 3786 GTCCGATCTGTACCCACTGACGGGGCACAGAGAGAGAGGAGTTCTCTCCAGAAAGAG 3845
 QY 1188 GluProSerThrLeuProAsnAsnValProAlaGlnGluGluIleGlyGluGluProGly 1207
 Db 3846 AGGCTCCAGCA---CTTCCAGTTTGTGTTCAGAGAAAGAACTAAGAAACATCAAG 3902
 QY 1208 ---ArgAspValLeuGluProThrGlnGlnGluLeuThrAlaAlaValProValLeu 1226
 Db 3903 ATGGAAGACACTTGAACATACAGATTAAGAGGTGACAGTGAAGATCTGTCCATTCTG 3962
 QY 1227 AlaLysThrGluValGlyGlnGluGluValAspTrpLeuAspGlyGlyLysValLys 1246
 Db 3963 TCAAGACTGAGGGGACTCAA-----GAGGCTGACCACTTCTGATGAGAAACCAAA 4016
 QY 1247 -----GluGlnGlnGluValPheValHisSerGly-----ProAsnSer 1259
 Db 4017 GACGTACCAATTTTGAAGGACCTGAGGGGTCTATAGACACAGGCAATACAGTCACTGG 4076
 QY 1260 GlnLysAlaAlaAspValThrTyraAspSerGluValMetGlyValAlaGlyCysGlnGlu 1279
 Db 4077 GAAAAAGTCACTGAAAGTTGCCCTTAAAGTGAAGGACAGAGCAAGAAAGCTGAATGTAAAAAG 4136
 QY 1280 LysGluSerThrGluValGlnSer-----LeuSerLeuGluGluGlyGlu 1294
 Db 4137 GATGATGCTCTTGAACCTGACAGTCAACCTTAAGTCTCTCAATCCCGGTGAGAGAGAG 4196
 QY 1295 MetGluThrAspValGluLysSerGluLysArgGluThrLysProGluGlnValSerGluGlu 1314
 Db 4197 ATGTAGATTCAAGTGAAGAGGAGAAACAGAGACAGAGCAACCCAGTGAATGAAGAG 4256
 QY 1315 Gly---GluGlnGluThrAlaAlaProGluHisGluGlyThrTyraGlyLysProValLeu 1333
 Db 4257 AAGCTTGAACACAGAAACAGCTGTACCTGATCTGAAGAGGTCAATACAGTCACTCCAG 4316
 QY 1334 ThrLeuAspMetProSerSerGluArgGlyLysAlaLeuGlySerLeuGlyGlySer--- 1352
 Db 4317 ACAGTGAATGTCCCATATAGATGGGCAAGAGAGATCAGCAGTTTGAAAGAGAGCCCT 4376
 QY 1353 ProSerLeuProAspGluAspLysAlaGlyCysIleGluValGlnValGlnSerLeuAsp 1372
 Db 4377 CTTCCCTGCTGATGATCAAGAGAGGCAATGACCAAAATTCAGAGTCTCAAGCTCTGAG 4436

Qy	1373	ThrThrValThrGlnThrAlaGluAlaValGluLysValIle-----GluThrValVal	1390
Db	4437	GCATCTACTTACACAGCGCTGCAGAGAGAAAGCTTTAGAGAAATGCCAAC	4496
Qy	1391	IleSerGluThrGlyGluSerProGluCysValGlyAlaHisLeuLeuProAlaGluLys	1410
Db	4497	ATTTAGAAACAGTGAACGTTGAGCGCTGCAGGTGCACATTTAGTTCTGGAAGAGAA	4556
Qy	1411	SerSerAlaThrGlyGlyHisTrpThrLeuGlnHisAlaGluAspThrValProLeuGly	1430
Db	4557	TCCTCTGAAAAAATGAAGACTTTTGGCGCTCATCCAGGGAAGATGCTGTGCCACAGG	4616
Qy	1431	ProGluSerGlnAlaGluSerIleProIleValThrProAlaProGluSerThrLeu	1450
Db	4617	CCGACGTGTGAGCAAAATGACACCATGATAGTATCTGCTACTACCAAGAAAGGCTTA	4676
Qy	1451	HisProAspLeuGlnGlyGluIleSerAlaSerGlnArgGluArgSerGluGluGluAsp	1470
Db	4677	AGTTCGACCTGGAAGAGAGAGAAACCATCATCTGAAGTGAAGTCAGATGAAGTCAT	4736
Qy	1471	LysProAspAlaGlyProAspAlaAspGlyLysGluSerThrAlaIleGluLysValLeu	1490
Db	4737	GAGCAGGTGCTTCCAGGAGGTC-----AAAGTGAGTGAGCAATTGAGGATTTA--	4787
Qy	1491	LysAlaGluProGlu-----IleLeuGluLeuGluSerLysSerAsnLysIleValLeu	1508
Db	4788	-----GAGCTGAAATGGGATTTTGGAACTTGAGACAAAAGCAGTAAACTTGTCCAA	4841
Qy	1509	AsnValIleGlnThrAlaValAspGlnPheAlaArgThr---GluThrAlaProGluThr	1527
Db	4842	AACATCATCAGACAGCGCTTGACCACTTGCTGCTACAGAGAAACAGCCAGCAATG	4901
Qy	1528	HisAlaTyrAspSerGlnThrGlnValProAlaCysArgLeuAspSerArgGluProAsn	1547
Db	4902	TGACGCTGCTGAGTTACAGACACAGCTCACGTGATAAAGCTGACAGCCAGCGCTGA	4961
Qy	1548	ArgCysTrpThrLys-----MetLysAspAla-LysMetLysHisProVa	1562
Db	4962	CAGGAACCGAGAAAGAGGAGAGAACCTCAGGCCTCTGCACAGGATGAACACCAAT	5021
Qy	1562	lProGlnProArgGlu-----AspLeuGlnValLeuThrValLe	1575
Db	5022	ACTTCAGCAAGAGGAGTACAGTCAACCGCAGTGGGACAGCAATCTTGATATTTC	5081
Qy	1575	uGluAlaTrpAlaGlnProArgLysCysLeuProArgLeuGlnLeuLysAlaProVal	1594
Db	5082	AAAGACATGAGTGAAGCTTCAGAAAGACCATGACTGTTGAGGTAGAGGTTCCACTG	5139
RESULT 11			
HSU81607			
LOCUS	HSU81607	6608 bp	linear
DEFINITION	Homo sapiens		
ACCESSION	U81607		
VERSION	U81607.1		
KEYWORDS			
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 6608)		
JOURNAL	Nauert, J.B., Klauck, T.M., Langeberg, L.K. and Scott, J.D.		
PUBMED	Gravin, an autoantigen recognized by serum from myasthenia gravis		
REFERENCE	patients, is a kinase scaffold protein		
AUTHORS	Curr. Biol. 7 (1), 52-62 (1997)		
TITLE	97153077		
JOURNAL	2000000		
PUBMED	2 (bases 1 to 6608)		
REFERENCE	Nauert, J.B., Klauck, T.K., Langeberg, L.K. and Scott, J.D.		
AUTHORS	Direct Submission		
TITLE	Submitted (10-DEC-1996) Biochemistry, OHSU, 3181 SW Sam Jackson		
JOURNAL	Parway, Portland, OR 97201, USA		
COMMENT	On Jun 26, 1997 this sequence version replaced gi:1850362.		
FEATURES	Location/Qualifiers		

source	1. .6608		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
CDS	192..5537		
	/note="kinase scaffold protein"		
	/codon_start=1		
	/product="gravin"		
	/protein_id="AAC51366.1"		
	/gi_xref="GI:2218077"		
	/translation="MGAGSSTBQRPEOPESSTPABPEPSGGSPAAEPADPTTADP		
	ATAASDPATKLQKNGQLSTINGVAEQDELSIQEGDLNGOKGALNGQALNGQALNQEEREVE		
	ITVEVQRDSERDSDKENATKSAVVVDHIDTDQOENRNIEQIPSSSENLBETQ		
	PTEQANDIGFKVFKFVKFTVKRDKPTDPTVOLLTVKKDQEGGAAGAHQDHP		
	LGAGAAKSESPKSTQETKPEETLRQSHAEISPPAESQVAGEKEKQEKQEP		
	SKSASPTSPVSTGTFKFTQCAWGRKTKPRKPKDEDEVAEKKEQEPKPV		
	DTEBQKAEVASEKLTAQEAHPAESAHPRLSAEYKVELPSEOVSGSQGPE		
	EKPAPLATEVDEKLEVEHQEVVAEVHVTVEERTEOKTEVEETAGSVPAELVGM		
	APQAEAPAKELVKLTCVSGEDTQADLSPDEKVLKPKPEGVVSEVMSSQERM		
	KVQGSPLKLFSTGLKLSGKKQKRGGEDESGEHTQVPADSPDSQEQKQESSA		
	SSPEPEITCLEGLAEVQDGEAEAGTSDEKKREGVTPWASFKMTVPKRVRR		
	PSKDEDELKVKATLSTESTASEMQEEMKSGVEPKPEPKRVKVDTSWDEALI		
	CVGSKGRARRSSSDEGGPKMGDHQKADKAGKDETDGILAGSOEHDPGGS		
	SSPEOAGSPTEGEGVSWESFKELATPRKSKSKLEEKSEDSIAGSVEHSTPDEPG		
	KZESVWSIKFIPGRKKRPDGKQQAPEVDAGTGANEDSDVPAVFLSYDAVER		
	EKMEAQKQAGAEPEQKAATEVSKELSESQVHMAAAVADGTRAATIIIEERSPWIS		
	ASVTFPLEQVEAAALTEVLEREVIIEEPTVTTEPLPENREARGDTPVSEAEITP		
	EAVTAETAGPLGSEGTASAABETEMVASVOLTSDPTTETATPVQVEEGLVPD		
	IEQERRTOEVLQVAEVKESQLPGTGDPEDVLPQVORAEARPEQAEASGLAKE		
	TDVVLVDAQEAETPFQGVVGTTPSPFEKAPQVTESELSLVTTCQETLAGV		
	KSQEMVQAIPPDSEVETDSTGTFVADFDAGTGTQKDEIWEIHENEVASGTQ		
	SGGTAEAVPAQKRPAPSPFQETKQSKMBDTLEHTDKVESVETVLSKTEG		
	TOEAQYADPKTKVPFEGLESIDTGTIVSRKVEVALKGEETEEAECKDDALE		
	LQSHAKSPSPVEREMVQVREKTEAETHVNEKLEHETATVTVSEEVSKOLLQTVN		
	VPDIQAEVSVLEGSPPCLQGEAACTKIOVQSSSEASFTLTAAREEKVLGETANI		
	LETGTLSPAGAHVLEKSKSEKNEDEFAHPGEDAVPTGPDCAKSTPVIYANTYKG		
	LSSDLEGKRTSLKAKSEVDEQVACQVKSVAIDDEPENGLIELETKSKLVONI		
	IQTAQVFRTEETATEMLTSELQOAHVKAQSDAQGETEKEBEPQSAQDEPTPI		
	TSKASESTAVQGAHSIKSDMSAEKMTVEVGTGTVNDQVLEEVLPSEEEGGG		
	AGTKSPVEDDGHALLAERIISKLVPKDEKGDVDDPENQALADTDAAGLTAKES		
	PDTNGPKKEKEDAEVLEQKGVHSESDKAITTPQAQBELQKQERESAKSELTES"		
BASE COUNT	2091 a 1418 c 1827 g 1272 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	1.85e-145	Length:	6608
Score:	4300.50	Matches:	967
Percent Similarity:	68.81%	Conservative:	189
Best Local Similarity:	57.56%	Mismatches:	408
Query Match:	53.27%	Indels:	116
DB:	9	Gaps:	36
US-09-902-432-4 (1-1596) x HSU81607 (1-6608)			
Qy	1	MetGlyAlaGlySerSerThrGluGlnArgSerProGluGln---ProAlaGlySerAsp	19
Db	192	ATGGGCGCGGAGCTCCACCGAGCAGCGCCGCGAGCAGCGCCGCGAGGAGTCC	251
Qy	20	ThrProSerGluLeuValLeuSerGlyHisGlyProAlaAlaGluAlaSer---GlyAla	38
Db	252	ACCGCGGCTGAGCCGCGAGCCCGCGCGCGCCCTCGCGCGAGCGCGCCAGACACC	311
Qy	39	AlaGlyAspProAlaAspAla-----AspProAlaThrLysLeuProGlnLysAsnGly	56
Db	312	ACCGCGGACCCCGCCATCGCTCGCTCGGACCCCGCCACCAAGCTCTCTACAGAGATGGT	371
Qy	57	GlnLeuSerSerValAsnGlyValAlaGluGlnGlyAspValHisValGlnGluAsn	76
Db	372	CAGCTGTCCACCATCAATGGCGTAGCTGAGCAAGATGAGCTCAGCTCCAGGAGGTGAC	431
Qy	77	GlnGluGlyGln-----	82
Db	432	CTAAATGGCCAGAAAGAGGAGCCCTGAAACGTCGTAAGGAGCCCTAAACAGCCAGGAGGA	491

QY 83 GIValValAspGluAspValGlyGlnArgGluSerGluAspValArgGluLysAspArg 102
 DB |||||.....
 DB 492 GAAGTCATTGTCA CGAGGTTGACAGAGACCTTGAAAGATGTGACCAAGAGACTCC 551
 QY 103 ValGluMetAlaAlaAsnSerThrAlaValGluAspIleThrLysAspGlyGlnGlu 122
 DB |||||.....
 DB 552 GATTAAAGATGGCTACTAAGTCAGCGGTTGTTACAGACTCAAGATGATGGGCGAG 611
 QY 123 GlnThrSerGluIleIleGluGlnIleProAlaSerGluLysAsnValGluGluMetVal 142
 DB |||||.....
 DB 612 GAGAACCGGAAT--ATCCAGACGATTCCTCTTCAGAAAGCAATTAAAGCGCTAAC 668
 QY 143 GlnProAlaGluSerGlnIleAsnAspValGlyPheLysLysValPheLysPheValGly 162
 DB |||||.....
 DB 669 CAACCACTGAGTCCAGGCTTAATGATTTGAGTTTAAGAGGTGTTTAAGTTTGTGGC 728
 QY 163 PheLysPheThrValLysLysAspLysAsnGluLysSerAspThrValGluLeuLeuThr 182
 DB |||||.....
 DB 729 TTTAAATTCACTCTGAAAAAGATTAAGACAGAAAGCCGACCTGACAGTCACTCACT 788
 QY 183 ValLysLysAspGluGlyGlyGlyAlaGluAlaSerValGlyAlaGlyAspHisGlnGlu 202
 DB |||||.....
 DB 789 GTGAAGAAAGATCAAGGAGGAGCA-----GCAAGGCGCTGCGACCAACGAGAG 839
 QY 203 ProSerValGluThrAlaValGlyGluSerAlaSerLysGluSerGluLeuLysGlnSer 222
 DB |||||.....
 DB 840 CCCAGCTT-----GGGCGTGGAGAGAGCAATCCAAAGAAAGCCAAACCAATCT 893
 QY 223 ThrGluLysGlnGluGlyThrLeuLysGlnGluGlnSerSerThrGluIleProLeuGln 242
 DB |||||.....
 DB 894 ACAGAGAAACCCGAAGAGACCTGAAAGCGTGAAGCAAGCAAGCAACCAATTTCTCCCA 953
 QY 243 AlaGluSerAspGlnIleAlaGluGluGluAlaLysAspGluGlyGluGlnGlu 262
 DB |||||.....
 DB 954 GCCGATCTGGCCCAAGCAAGT--GAGGAATGCCAAAGAGAGAGAAAGAAACAAGAA 1010
 QY 263 LysGluProThrLysSerProGluSerProSerSerProValAsnSerGluThrThrSer 282
 DB |||||.....
 DB 1011 AAGAACTCAGCAAGTCTGAGATCTCCGACTGATCCGTTGACAGTGAAGAAAGATCA 1070
 QY 283 SerPheLysLysPhePheThrHisGlyTyrAlaGlyTyrArgLysLysThrSerPheLys 302
 DB |||||.....
 DB 1071 ACCTTAAAAAATTTCTCACTCAAGTTGGCGCGCTGGCGCAAAAAGCAATTTCAAG 1130
 QY 303 LysSerLysGluAspAspLeuGluThrAlaGluLysArgLysGluGlnGluAlaGluLys 322
 DB |||||.....
 DB 1131 AAGCCAGAGAGATGAGTGAAGTGAAGCTTCAGAGAAAGAAAGAAAGCAAGCGCAAGAAA 1190
 QY 323 ValAspGluGluGluLysGluLysThrGluProAlaSerGluGlu----- 337
 DB |||||.....
 DB 1191 GTTAGACACAGAAAGACGGAAGGACAGAGTTGCTCCGAGAACTGACCGGCTCCGAG 1250
 QY 338 -----GlnGluProAlaGluAspThrAspGlnAlaArgLeuSerAlaAspTyr 353
 DB |||||.....
 DB 1251 CAAGCCACCCACAGAGACCGGCGAGAAAGGCCACAGAGCCCGCTTATAGTGAATAT 1310
 QY 354 GlnLysValGluLeuProLeuGluAspGlnValGlyAspLeuGluAlaSerSerGluGlu 373
 DB |||||.....
 DB 1311 GAGAAAGTTGAGTCCCTCAGAGAGCAAGTCAAGTGGCTCGCAGGAGACTTCTGAAGAG 1370
 QY 374 LysCysAlaProLeuAlaThrGluValPheAspGluLysMetGluAlaHisGln---Glu 392
 DB |||||.....
 DB 1371 AAACCTGCTCCGTTGGCGAGAGAGTGTGAGAGAAATAGAAAGTCCCAAGAGAGAG 1430
 QY 393 ValValAlaGluValHisValSerThrValGluLysThrGluGluGlnGlnGlyGly 412
 DB |||||.....
 DB 1431 GTTGTGGCGAAATCCAGCTCAGACCGTGGAGAGAGAAACCAAGAGCAG----- 1481
 QY 413 GlyGluAlaGluGlyGlyValValGluGlyThrGlyGluSerLeuProGluLys 432
 DB |||||.....
 DB 1482 -----AAACGAGAGTGAAGAAACAGAGGCTGTGCGCAGGTGAAGAA 1526
 QY 433 LeuAlaGluProGlnGluValProGlnGluAlaGluProAlaGluLeuMetLysSer 452

DB 1527 TTGGTTGAATGATGCAAGACCTCAGAGACCGAACCTGCAAGAGAGCTGTAGACTC 1586
 QY ArgGluMetCysValSerGlyLysAspHisThrGluLeuThrAspLeuSerProGluGlu 472
 DB |||||.....
 DB 1587 AAAGAAAGTGTGTTCGAGAGAGACCTTACAGAGGAGCTGACCTCAGTCTGATAG 1646
 QY 473 LysThrLeuProLysHisProGluGlyIleValSerGluValGluMetLeuSerGln 492
 DB |||||.....
 DB 1647 AAGGTGCTGCCAAACCCCGAGAGCGTGTGAGTGAAGTGAAGAAATGCTTCACTCAG 1706
 QY 493 GluArgIleLysValGlnGlySerProLeuLysLysLeuPheSerSerSerGlyLeuLys 512
 DB |||||.....
 DB 1707 GAGAGAAAGAAAGTCCAGAGAAAGTCCATAAAGAAAGCTTTTACAGACAGTGGCTTAA 1766
 QY 513 LysLeuSerGlyLysLysGlnLysGlyLysArgGlyGlyGlyAspGluGluProGly 532
 DB |||||.....
 DB 1767 AAGCTTCTGGAAGAAACAGAAAGGAAAGA--GGAGAGAGAGACGAGAAATCAGGG 1823
 QY 533 GlnTyrGlnHisIleHisThrGluSerProGluSerAlaAspGluGlnLysGlyGluSer 552
 DB |||||.....
 DB 1824 GAGCACTCAGGTTCCAGCCGATTTCCGAGACGCCAGAGAGCAAAAGGCGAGAGC 1883
 QY 553 SerAlaSerSerProGluGluProGluGluThrThrCysLeuGluLysGlyProLeuGlu 572
 DB |||||.....
 DB 1884 TCTGCTCATCCCTGAGAGACCCGAGAGATCACTGCTCGAAAGAGCTTACCCGAG 1943
 QY 573 AlaProGlnAspGlyGluAlaGluGlyThrThrSerAspGlyGlyLysArgGlu 592
 DB |||||.....
 DB 1944 GTGCAGCAGAGATGGGAGAGTGAAGAAAGAGACTCTTCCGATGGAGAAAGAAAGAA 2003
 QY 593 GlyIleThrProThrAlaSerPheLysLysMetValThrProLysLysArgValArgArg 612
 DB |||||.....
 DB 2004 GGTGTCACTCCCTGGGCAATCAATCAAAAGATGGAGAGCCCAAGAACGGTGTAGAGG 2063
 QY 613 ProSerGluSerAspLysGluGluLeuGluLysValLysSerAlaThrLeuSerSer 632
 DB |||||.....
 DB 2064 CTTTGGAAAGATTAAGAAAGATAGCTGACAGAAAGTCAAGAGGCTTACTTGTCTTCC 2123
 QY 633 ThrAspSerThrValSerGluMetGlnAspGluValLysThrValGlyGluGluLys 652
 DB |||||.....
 DB 2124 ACCGAGAGACAGCTCTGAAATGCAGAAAGAAAGAGAGCGCTGGAAGACCAAG 2183
 QY 653 ProGluGluProLysArgArgValAspThrSerValSerTPGluAlaLeuIleCysVal 672
 DB |||||.....
 DB 2184 CCGAAGAAACCAAGCGAAGGTGATACCTCAATCTTGGAGAGCTTTAATTTGTGTG 2243
 QY 673 GlySerSerLysLysArgAlaArgLysAlaSerSerSerAspAspGluGlyProArg 692
 DB |||||.....
 DB 2244 GATCATCCAAAGAAAGCAAGAGAGAGAGTCTCTTCTGATGAGAGAGGGGACCAAA 2303
 QY 693 ThrLeuGlyLysAspSerHisArgAlaGluGluAlaSerLysAspLysGluAlaGlyThr 712
 DB |||||.....
 DB 2304 GCAATGGAGAGAGACCAACAGAAAGCTGATGAGCGCGGAGAAAGAACAGAGCGGACA 2363
 QY 713 AspAlaValProAlaSerThrGlnGluGlnAspGlnAlaGlnGlySerSerProGlu 732
 DB |||||.....
 DB 2364 GACGGATCTCTTGTGTTCCCAAGAACATGATCCAGGGCGAGGAAAGTTCTCCCGAG 2423
 QY 733 ProAlaGlySerProSerGluGlyGlnGlyValSerThrTPGluSerPheLysArgLeu 752
 DB |||||.....
 DB 2424 CAAGCTGAAAGCCCTTACCGAAGGAGGCGTTTCACTGGAGATCTTTAAAGGTTA 2483
 QY 753 ValThrProArgLysLysSerLysSerLysLeuGluGlyLysAlaGluAsp----- 769
 DB |||||.....
 DB 2484 GTCAAGCCCAAGAAATAATCAAGTCCAAGCTGAAGAGAAAGAGAAAGACTCCATAGCT 2543
 QY 770 ---SerSerValGluGlnLeuSerThrGluIleGluProSerArgGluGluSerTPVal 788
 DB |||||.....
 DB 2544 GGTCTGTGTGTGAACATTTCCACTCCAGACCTGAACCCGCTAAAGAAAGATCTCGGGTCC 2603
 QY 789 SerIleLysPheIleProGlyArgArgLysLysArgAlaAspGlyLysGlnGluGln 808

Db	2604	TCATCATCAAGAAAGTTTATTCCTCGCAGCAAGCAAGAAAGGCCAGATGGGAAACAAGAACAA	2666
Qy	809	AlaThrValGluAspSerGlyProValGluIleAsnGluAspAspProAsnValProAla	828
Db	2664	CCCCCTGTTGAAGACGACAGGCCCAACAGGGGCCACAGAGATGACTCTGTATGTCCCGGCC	2723
Qy	829	ValValProIeuSerGluTyrAenAlaValGluArgGluLysMetGlu-----Ala	845
Db	2724	GTGGTCCCTCTGTCTGAGTATGATGCTGTAGAAAGGGAGAAAAATGGAGCCACAGCAAGCC	2783
Qy	846	GlnGlyAenThrGluLeuProGlnLeuLeuGlyAlaValTyrValSerGluGluLeuSer	865
Db	2784	CAAAAGCCGACAGACAGCCCGAGCAGAGACGACGACCTAGAGGTCTCCAGGAGCTCAGC	2843
Qy	866	LysThrLeuValHisThrValSerValAlaValIleAspGlyThrArgAlaValThrSer	885
Db	2844	GAGAGTCAGGTTCATATGATGGCAGCAGCTGTCTGACGGACGACGGGCGAGCTACCATTT	2903
Qy	886	ValGluGluArgSerProSerTrpIleSerAlaSerValThrGluProLeuGluHisThr	905
Db	2904	ATTGAAGAAAGGTCCTCTTCTTGATATCTGCTTCAGTCACAGAACCTCTTTGAACAAGTA	2963
Qy	906	AlaGlyGluAlaMetProProValGluGluValThrGluLysAspIleAla---Glu	924
Db	2964	GAAGCTCAAGCCGCACCTGTTAACTGAGGAGGTATTGGAAGAAGAAAGTAATTTGCAAGAA	3023
Qy	925	GluThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHisAspAspMetVal	944
Db	3024	GAACCCCCACGGTTACTGAACTCTTCGCAGAGAACAGAGAGCGCCGGGGCGACACCGGTC	3083
Qy	945	ThrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeu	964
Db	3084	GTTAGTGAGGCGGAATTGACCCCGAAGCTGTGACAGCTGCAGAAACTGCAGGGCCATTG	3143
Qy	965	ArgThrGluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSerAla	984
Db	3144	GGTTCGGAAGAGCAACCGCAAGCATCTGCTGCTGAGAGACACACAGAAATGGTGTCAACA	3203
Qy	985	ValSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGluVal	1004
Db	3204	GTCTCCAGTTAACCGACTCCCAAGACACCAAGAGGAGGCCACTCCGGTGCAGAGGTG	3263
Qy	1005	GluSerGlyValLeuAspThrGluGluGluArgGlnThrGlnAlaIleLeuGlnAla	1024
Db	3264	GAAGGTGGGTACCTGACATAGAAGACGAAGAGAGCGGACTCAAGAGGTCTCTCCAGCA	3323
Qy	1025	ValAlaAspLysValLysGluGluSerGlnValProAlaThr-----	1038
Db	3324	GTGCGCAAAAAGTGAAGAGGAATCCCACTGCTGCGCAGCGGTGGGCCAGAGAGATGTG	3383
Qy	1039	--GlnThrValGlnArgThrGlySerLysAlaLeuGluLysValGluGluValGluGlu	1057
Db	3384	CTTCAGCTGTGCAGAGAGCA-----GAGCGAGAAGACCAAGAGAG	3425
Qy	1058	AspSerGluValLeuAlaSerGluLysGluLysAspValMetProLysGlyProValGln	1077
Db	3426	CAGCTGAAGCGTCGGTCTGAAGAAGAGACCGGATGTAGTGTTCGAAGTAGTAGTCTCAG	3485
Qy	1078	GluAlaGlyAlaGluHisLeuAlaGlnGlySerGluThrGlyGlnAlaThrProGluSer	1097
Db	3486	GAGGCAAAAACCTGAGCCCTTTTACCAAGGGGAAGGTGGTGGGCGACACCAACCCCAAGAA	3545
Qy	1098	LeuGlu---ValProGluValThrAlaAspValAspHis-----ValAlaThrCys	1113
Db	3546	TTTGAAAAGAGCTCCTCAAGTCAAGAGACGATAGAGTCCAGTGAGCTTGTGAACCACTGT	3605
Qy	1114	GlnVal-----IleLysLeuGlnGlnLeu---MetGluGlnAlaValAla	1127
Db	3606	CAAGCCGAACCTTAGCTGGGGTAAATACAGGAGATGGTGTATGGACAGAGCTATCCCC	3665
Qy	1128	ProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrProLeuAlaAsp	1147
Db	3666	CCTGACTCGGTGAACCCCTACAGACGTGAGACTGTATGGAAGCAACCCCGTGGAGCCGAC	3725

Qy	1148	SerAspThrAlaAspGlyThrGlnGlnAspGluThrIleAspSerGlnAspSerLysAla	1167
Db	3726	TTTCACGACACAGGCACAACCCAGAAAGACAGAGATTGTGGAAATCCATGAGGAGAAATCGAG	3785
Qy	1168	ThrAlaAlaValArgGlnSerGlnValThrGluGluGluAlaAlaThrAlaGlnLysGlu	1187
Db	3786	GTCGCATCTGGTACCAGTCAGGGGCGCACAGAAAGCAGAGGCGAGTTCCTGCACAGAAAGAG	3845
Qy	1188	GluProSerThrLeuProAsnAsnValProAlaGlnGlnGluHisGlyGluGluProGly	1207
Db	3846	AGGCCTCCAGCA--CCTTCCAGTTTGTGTTCTCCAGGAAGAACTAAAGAAACAATCAAG	3902
Qy	1208	--ArgAspValLeuGluProThrGlnGlnGluLeuLeuThrAlaAlaAlaValProValLeu	1226
Db	3903	ATGGAAGACACATCTTAGAGCATACAGATAAAGAGGTGTCCAGTGGAAACTGTATTCATTCG	3962
Qy	1227	AlaLysThrGluValGlyGlnGluGlyGluValAspTrpLeuAspGlyGluLysValLys	1246
Db	3963	TCAAAGACTCAGGGGACTCAA-----GAGGCTGACCAATGATGCTGATGAGAAACCAAA	4016
Qy	1247	-----GluGluGlnGluValPheValHisSerGly-----ProAsnSer	1259
Db	4017	GACGTACCATTTTTTCNAGGACTTGAGGGGTCTATAGACACAGGATACACAGTCAGTCGG	4076
Qy	1260	GlnLysAlaAlaAspValThrTyAspSerGluValMetGlyValAlaGlyCysGlnGlu	1279
Db	4077	GAAAGGTCACTGAAGTTGCCCTTAAAGGTGAAGGCAGAGCAAGAGCTGAATGTAAAGAG	4136
Qy	1280	LysGluSerThrGluValGlnSer-----LeuSerLeuGluGluGlyGlu	1294
Db	4137	GATGATGCTCTTGAACATGCAGAGTACGCTAAAGTCTCTCCATCCCGCTGGAGAGAGAG	4196
Qy	1295	MetGluThrAspValGluLysGluLysArgGluThrLysProGluGlnValSerGluGlu	1314
Db	4197	ATGTGATGTCAGTCAAGTCAAGAGGGGAGAAAACAGAAAGCAGACCAACCATGTGAATGAAGAG	4256
Qy	1315	Gly----GluGlnGluThrAlaAlaProGluHisGluGlyThrTyGlyLysProValLeu	1333
Db	4257	AAGCTTGAGCAGCAAAACAGCTGTACCGTATCTGAAGAGGTCAGTAAGCAGCTCCTCCAG	4316
Qy	1334	ThrLeuAspMetProSerSerGluArgGlyLysAlaLeuGlySerLeuGlyGlySer---	1352
Db	4317	ACAGTGAATGTGCCCATCATAGATGGGCAAAAGGAAGTCAGCAGTTTGGAAAGGAGCCCT	4376
Qy	1353	ProSerLeuProAspGlnAspLysAlaGlyCysIleGluValGlnValGlnSerLeuAsp	1372
Db	4377	CCTCCCTCCCTAGGTCAAGAGAGGCAGATATGCACCAAAATTCAGTTTCAGAGCTCTGAG	4436
Qy	1373	ThrThrValThrGlnThrAlaGluAlaValGluLysValIle-----GluThrValVal	1390
Db	4437	GCATCATTTCACTTACACGGCGCTGCAGAGGAGGAAAAGGTCTTAGGAGAAACTGCCAAC	4496
Qy	1391	IleSerGluThrGlyGluSerProGluCysValGlyAlaHisLeuLeuProAlaGluLys	1410
Db	4497	ATTTTAGAAACAGGTGAACGTTGGAGCCCTGCAGGTGCACATTTAGTTCTCGAAGAGAGAA	4556
Qy	1411	SerSerAlaThrGlyGlyHisTrpThrLeuGlnHisAlaGluAspThrValProLeuGly	1430
Db	4557	TCCCTCTGAAAAAATAAGAACCTTTCCGCTCATCCAGGGGAGAGATGCTGTGCCACAGGG	4616
Qy	1431	ProGluSerGlnAlaGluSerIleProIleIleValThrProAlaProGluSerThrLeu	1450
Db	4617	CCCGACTGTCCAGCAAAATCGACACAGTGAATGATATCTGCTATCCACAGAAAGGCTTA	4676
Qy	1451	HisProAspLeuGlnGlyGluIleSerAlaSerGlnArgGluArgSerGluGluGluAsp	1470
Db	4677	AGTTCCGACCTGGNAGGAGAGAAACCATCACTCAAGTGAAGTCAGATGAAGTCGAT	4736
Qy	1471	LysProAspAlaGlyProAspAlaAspGlyLysGluSerThrAlaIleGluLysValLeu	1490
Db	4737	GAGCAGGTGCTGTCAGCAGAGGTC-----AAAGTGAGTGTAGCAATTCAGCATTTTA	4787

QY 433 LeuAlaGluProGlnGluValProGlnGluAlaGluProAlaGluGluLeuMetLysSer 452
DB 1527 TTGTTTGAATGGATGAGAACCTTCAGAAAGCCGAACCTGCAAGGAGCTGGTGAAGCTC 1586
QY 453 ArgGluMetCysValSerGlyGlyAspHisThrGlnLeuThrAspLeuSerProGluGlu 472
DB 1587 AAAGAAACGTGTGTTTCCGAGAGGACCTACACAGGAGCTGACCTCAGTCTCTGATGAG 1646
QY 473 LysThrLeuProLysHisProGluGlyLeValSerGluValGluMetLeuSerSerGln 492
DB 1647 AAGGTGCTGTCCAAACCCCGAAGGCGTTGTGAGTGAGGTGGAATGCTGTATCATCAG 1706
QY 493 GluArgIleValGlnGlnGlySerProLeuLysLysLeuPheSerSerSerGlyLeuLys 512
DB 1707 GAGAAATGAGGTGCGAGGAGTCCACTAAAGAAAGCTTTTACCACTGCGCTTAAAA 1766
QY 513 LysLeuSerGlyLysLysGlnLysGlyLysArgGlyGlyGlyLysAspGluGluProGly 532
DB 1767 AAGCTTCTGGAAGAAACAGAAAGGGAAGA---GGAGGAGGAGACGAGGAATCAGGG 1823
QY 533 GluThrGlnHisIleHisThrGluSerProGluSerAlaAspGluGlnLysGlyGluSer 552
DB 1824 GAGCACACTCAGGTTCACGCGCATTTCCGACACAGCAGGAGGAGCAAAAGGGCGAGAGC 1883
QY 553 SerAlaSerSerProGluGluProGluGluThrThrCysLeuGluLysGlyProLeuGlu 572
DB 1884 TCTGCCCTCATCCCTGAGGAGCCGAGAGATCATCGTGTCTGGAAGAGGGCTTAGCCGAG 1943
QY 573 AlaProGlnAspGlyGluAlaGluGluGlyThrThrSerAspGlyGluLysLysArgGlu 592
DB 1944 GTGCAGCAGATGGGGAAGCTGAAGAGGAGCTACTTCCGATGAGAGAGAAAAAGAGAA 2003
QY 593 GlyIleThrProThrAlaSerPheLysLysMetValThrProLysLysArgValArg 612
DB 2004 GGTGTACTCCTCGGCATCATTTCAAAAGATGTTGACGCCCAAGAACGCGTTAGACGG 2063
QY 613 ProSerGluSerAspLysGluGluGluLysValLysLysValLysSerAlaThrLeuSerSer 632
DB 2064 CCTTCGAAAGTGAATAAGAGATGAGCTGACAAAGTCAAGAGCGCTTACCTGTCTTCC 2123
QY 633 ThrAspSerThrValSerGluMetGlnAspGluValLysThrValGlyGluGluGlnLys 652
DB 2124 ACCGAGACACAGCTCTGAAATGCAAGAGAAATGAAGGAGCGTGGAGAGCCAAAG 2183
QY 653 ProGluGluProLysArgValAspThrSerValSerTrpGluAlaLeuIleCysVal 672
DB 2184 CCGGAACAACCAAGCCAAAGTGTGATCTCAGTATCTTGGGAAGCTTTAAATTTGTTG 2243
QY 673 GlySerSerLysLysArgAlaArgLysAlaSerSerSerAspAspGluGlyGlyProArg 692
DB 2244 GGATCATCCAAAGAAAGCAAGGAGAGGTCTCTTCTGATGAGGAGGCGGACCAAAA 2303
QY 693 ThrLeuGlyGlyAspSerHisArgAlaGluAlaSerLysAspLysGluAlaGlyThr 712
DB 2304 GCATGGGAGAGACCAACCAAGAAAGCTGATGAGCGCGGAAAGACAAAGAGACGGGACA 2363
QY 713 AspAlaValProAlaSerThrGlnGlnAspGlnAlaGlnGlySerSerSerProGlu 732
DB 2364 GACGGATCTTGTGTTCCCAAGACATGATCCAGGGCAGGGAAGTTCCTCCCGGAG 2423
QY 733 ProAlaGlySerProSerGluGlyGlyValSerThrTrpGluSerPheLysArgLeu 752
DB 2424 CAAGCTGGAAGCCCTACCGAAGGGGAGGGGCTTTCCACCTGGGAGTCAATTTAAAGGTTA 2483
QY 753 ValThrProArgLysLysSerLysSerLysLeuGluLysAlaGluAsp-----Val 769
DB 2484 GTACGCGCAAGAAAAAATCAAGTCCAGCTGGAAGAGAAAAAGCGAAGACTCCATAGCT 2543
QY 770 ---SerSerValGlnGlnLeuSerThrGluIleGluProSerArgGluGluSerTrpVal 788
DB 2544 GGGTCTGGTGTAGACATTTCCACTCCAGACACTGAACCCGGTAAAGAGAAATCTCGGTG 2603
QY 789 SerIleLysLysPheIleProGlyArgArgLysLysArgAlaAspGlyLysGlnGluGln 808

DB 2604 TCAATCAAGAAGTTTATTCTGACGAAGGAAGAAAGGCCAGATGGGAAACAGAAACA 2663
QY 809 AlaThrValGluAspSerGlyProValGluIleAsnGluAspAspProAsnValProAla 828
DB 2664 GCCCTGTTGAAGACGACGAGGCCCAACAGGGGCCCAACGAAGATGACTCTGTATGTC 2723
QY 829 ValValProLeuSerGluTyrAsnAlaValGluArgGluLysMetGlu-----Ala 845
DB 2724 GTGGTCCCTCTGTCTGATGATGCTGTAGAAAGGGAGAAAAATGGAGGCACAGCAAGCC 2783
QY 846 GlnGlyAsnThrGluLeuProGlnLeuGluValValTyrValSerGluGluLeuSer 865
DB 2784 CAAAAAGCCGACAGCAGCCGAGCAGAGGAGCCACTGAGGTGTCCAAGAGCTCAGC 2843
QY 866 LysThrLeuValHisThrValSerValAlaValIleAspGlyThrArgAlaValThrSer 885
DB 2844 GAGAGTCAGGTTCATATGATGGCAGCAGCTGTCTGCTGACGGAGGAGCGAGCTACCAT 2903
QY 886 ValGluGluArgSerProSerTrpIleSerAlaSerValThrGluProLeuGluHisThr 905
DB 2904 ATTGAAGAAAGGTCTCCTCTTGGATATCTGCTTTCAGTGACAGAACCTCTTGAACAA 2963
QY 906 AlaGlyGluAlaMetProProValGluGluValThrGluLysAspIleIleAla---Glu 924
DB 2964 GAAGCTGAAGCCGACCTGTTAACTGAGGAGGTATTGGAAGAGAGAAATTTGCAGAA 3023
QY 925 GluThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHisAspAspMetVal 944
DB 3024 GAACCCCCACGGTTACTGAACCTTCCAGAGAACAGAGAGGCCCGGGCGACACGGTC 3083
QY 945 ThrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeu 964
DB 3084 GTTAGTGAGCGGAAATTGACCCCGAAGCTGTGACAGCTGCAGAAACTGCAGGCCCAT 3143
QY 965 ArgThrGluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSerAla 984
DB 3144 GGTTCGGAAGAGAAACCGAAGCATCTGCTGCTGAAGAGACCACAGAAATGGTGTGAC 3203
QY 985 ValSerGlnLeuThrAspSerProAspThrThrGluAlaThrProValGlnGluVal 1004
DB 3204 GTCTCCAGTTAACCCACTCCCGACACACACAGAGAGGCCACTCGGTGCGAGGAGTG 3263
QY 1005 GluSerGlyValLeuAspThrGluGluGluArgGlnThrGlnAlaIleLeuGlnAla 1024
DB 3264 GAAGTGGCGTACTGACATAGAAGAGCAAGAGAGCGGACTCAAGAGGTCCTCCAGGCA 3323
QY 1025 ValAlaAspLysValLysGluGluSerGlnValProAlaThr-----Val 1038
DB 3324 GTGCGCAAAAAAGTGAAGAGGAATCCAGCTGCTGGCACCGGTGGGCCAGAGATGTG 3383
QY 1039 ---GlnThrValGlnArgThrGlySerLysAlaLeuGluLysValGluGluValGlu 1057
DB 3384 CTTACGCTGTGCAGAGACA-----GAGCAGANAAGACCAGAAAG 3425
QY 1058 AspSerGluValLeuAlaSerGluLysGluLysAspValMetProLysGlyProValGln 1077
DB 3426 CAGCTGAAGCGTCGGTCTGAAGAAAGAGACGCGATGTAGTGTGAAAGTAGATGCTCAG 3485
QY 1078 GluAlaGlyAlaGluHisLeuAlaGlnGlySerGluThrGlyGlnAlaThrProGluSer 1097
DB 3486 GAGGCAAAAACTGAGCCTTTTACACAAGGGGAAGGTGTGTGGGGCAGACCACCCAGAA 3545
QY 1098 LeuGlu---ValProGluValThrAlaAspValAspHis-----ValAlaThrCys 1113
DB 3546 TTTGAAAAAGCTCCTCAAGTCAAGAGAGCATAGATCCAGTGAGCTTTGAACCATCTGT 3605
QY 1114 GlnVal-----IleLysLeuGlnGlnLeu---MetGluGlnAlaValAla 1127
DB 3606 CAAGCCGAACCTTAGTGGGTAAATATCACAGAGATGTTGATGGAAACAGGCTATCCCC 3665
QY 1128 ProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrProLeuAlaAsp 1147


```

Db 3666 CCTGACTGGTGGAAACCCCTTACAGACAGTGAAGTGAAGACCCCTTACCCAG 3725
Oy 1148 SerApThrlaAspGlyThrGlnInaSpGluThrlaAspSerGlnaSerIysala 1167
Db 3726 TTGTGCGACGACGACAAACCCAGAAAGCAGATTGTGGAAATCCATGAGAGATGAG 3785
Oy 1168 ThrAlaIaValArgGlnSerGlnValThrGluGluGluAlaIaThrAlaGlnIysGlu 1187
Db 3786 GTGCATCTGTGA---CGATGCGGGGCAACGAGAGAGAGAGCTTCTGCGACAGAAAGG 3842
Oy 1188 GluProSerThrlaProAsnAsnValProAlaGlnGluGluIHisGlyGluGluProGly 1207
Db 3843 AGCCTCCAGCA---CCTTCCAGTTTGTGTTCCAGAGAAAGAACTAAAGACATCAAG 3899
Oy 1208 --ArgAspValLeuGluProThrGlnGlnGluLeuThrlaAlaIaValProValLeu 1226
Db 3900 ATGMAAGACACTGAGATACAGATAAAGAGGTGTGCTAGTGAAGCTATATCCATTCTG 3959
Oy 1227 AlaIysThrGluValGlyGlnGluGluValaIaAspTrpLeuAspGlyGluIysValIys 1246
Db 3960 TCAAGACTGAGGGGAGCTCA---GAGGCTGACACAGTATGCTGATGAGAAACCAAA 4013
Oy 1247 -----GluGluGlnGluValPheValHisSerGly-----ProAsnSer 1259
Db 4014 GACGTACCATTTTTCAGAGAGCTTGAAGGCTGTATAGACACAGGCATTAACAGTCACTGCG 4073
Oy 1260 GlnIysAlaIaAspValThrTyraSerSerGluValMetGlyValaIaGlyCysGlnGlu 1279
Db 4074 GAAAGGTCACGAAAGTTGCCCTTAAAGGTGAAGGACGAAAGAGCTGAATGTAAAG 4133
Oy 1280 LysGluSerThrlaGluGlnSer-----LeuSerLeuGluGluGluGlu 1294
Db 4134 GATGATGCTCTTGAACCTGAGAGTCACTGAAGTCTCTCCATCCCGCTGAGAGAGAG 4193
Oy 1295 MetGluThrlaAspValGluIysGlnIysArgGluThrlaYsProGluGlnIaIaSerGluGlu 1314
Db 4194 ATGTGATGTTCAAGTGAAGGAGGAGAAACAGAAACGACACCAACCATGTGAATGAAG 4253
Oy 1315 Gly---GluGlnGluThrlaAlaAspGluIHisGlyGluGluThrlaYsIysProValLeu 1333
Db 4254 AAGCTGAGCAGAAACAGCTTACCCGATGTAAGAGAGTCAAGTACAGAGCTCCCGCAG 4313
Oy 1334 ThrLeuAspMetProSerSerGluIaArgGlyValaIaLeuGlySerLeuGlyIysSer--- 1352
Db 4314 ACAGGAAATGTGCCATCATCATGAGGGGCAAGAGAGTCAAGTGTGGAAGAGAGCCCT 4373
Oy 1353 ProSerLeuProAspGlnaAspIysAlaGlyCysIleGluValGluValGlnSerLeuAsp 1372
Db 4374 CTTCCCTGCTAGTGTCAAGAGAGGAGATGATGACCAAAATTCAGAGCTCTGAG 4433
Oy 1373 ThrThrlaThrGlnThrlaIaGluIaValaIaGluIysValIle-----GluThrlaValaI 1390
Db 4434 GCATCATTCACCTTCAACGCGGCTGACAGAGAGAAAGTCTTAGAGAAACCTGCCAAC 4493
Oy 1391 IleSerGluThrlaGlyGluSerProGluCysValGlyAlaHisLeuLeuProAlaGluIys 1410
Db 4494 ATTTAGAAACAGGTGAACCTTGGAGCTGCAAGTGCACATTTAGTTCTGGAGAGAA 4553
Oy 1411 SerSerAlaThrGlyGlyHisIsthrThrlaLeuGlnHisAlaGluAspThrlaProLeuGly 1430
Db 4554 TCCTCTGAAAAAATGAACATTTGCCGCTCATCCAGGGGAAAGATGCTGTGCCACAGG 4613
Oy 1431 ProGluSerGlnaIaGluSerIleProIleIleValThrlaProAlaProGluSerThrla 1450
Db 4614 CCCGACTGTCAGGCAAAATCGACACCGATGATGATCTGCTACTCTCAAGAAAGCTTA 4673
Oy 1451 HisProAspLeuGlnGlyGluIleSerAlaSerGlnaArgGluIaArgSerGluGluIaAsp 1470
Db 4674 AGTTCGACCTCGAAGAGAGAAACCATCATGATGAAGTCAATGATGAAGTCGAT 4733
Oy 1471 LysProAspAlaGlyProAspAlaAspGlyIysGluSerThrlaIleGluIysValLeu 1490
Db 4734 GAGCAGGTTGCTTCCAGGAGGTC-----AAAGTAGTGTAGCAATTAAGGATTTTA--- 4784

```

```

Oy 1491 LysAlaGluProGlu-----IleLeuGluLeuGluIysSerIysAsnIysIleValLeu 1508
Db 4785 -----GAGCCTGAAAAATGGGATTTTGGACCTTGAACCAAAACAGATTAACCTGTCCA 4838
Oy 1509 AsnValIleGlnThrlaIaValaIaAspGlnPheAlaArgThr---GluThrlaProGluThr 1527
Db 4839 AACATCATTCAGACAGCCGTTGACAGTTTGTACGTACGAGAAAGAACCCAGCAAAATG 4898
Oy 1528 HisAlaTyraSerSerGlnThrlaValProAlaCysArgLeuAspSerArgGluProAsn 1547
Db 4899 TTGACGTGTGAGTTACAGACACAAGCTCAGTGATTAAGCTGACAGCCAGAGCGCTGA 4958
Oy 1548 ArgCysTrpThrIys-----MetLysAspAla-LysMetLysHisProva 1562
Db 4959 CAGMAACGAGAAAGAGAGAGAGAACCTCAGGCTCTGCGACGAGATGAAGAACCAAT 5018
Oy 1562 LProGlnProArgGlu-----AspLeuGlnValLeuThrlaValle 1575
Db 5019 ACTTCAGCCAAAGAGAGTCAAGTCAACCCAGTGGACAGACATTCGTGATTTTCC 5078
Oy 1575 uGluAlaTrpAlaGlnProArgIysCysLeuProArgLeuGlnLeuIysAlaProval 1594
Db 5079 AAAGCATGAGTGAAGCCTCAGAAAGAACCATGACTGTGAGTGAAGAGTTCCACTG 5136

RESULT 13
AF001504 6297 bp mRNA linear PRI 06-JUN-1997
LOCUS AF001504
DEFINITION Homo sapiens myasthenia gravis autoantigen gravin mRNA, partial cds.
ACCESSION AF001504
VERSION AF001504.1 GI:2160783
KEYWORDS .
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euteleia; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 6297)
AUTHORS Bowditch,R.D. and Ginsberg,M.H.
TITLE Sequence of gravin cDNA isolated from a human umbilical vein endothelial cell library
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 6297)
AUTHORS Bowditch,R.D. and Ginsberg,M.H.
TITLE Direct Submission
JOURNAL Submitted (29-APR-1997) Biochemistry, University of Oklahoma Health Sciences Center, PO Box 26901, BMSB 853, Oklahoma City, OK 73190, USA

FEATURES
source
Location/Qualifiers
1..6297
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="endothelial cell"
/tissue_type="umbilical vein"
<1..5526
/feature_type="kinase scaffold protein"
/codon_start=1
/product="myasthenia gravis autoantigen gravin"
/protein_id="AAB58938.1"
/db_xref="GI:2160784"
/translation="DPAIASDPATKLLQKNQGLSTINGAEDLSLQSGDLNGQK
ALNQGALNQEEBEVIVTEVGQSDSDVERSDKMATKSAVAVHITDQGETPE
IIEQIPSEBNLELTPTESQANDIDFKVFKFVGKFKLVKDKQKTEKPDVQLTVK
KDBEGGAAGADHDDPSLGAEGAESKSEPKOSKPEETLKREOSAEIISPAESGQ
AVBECKREGEKQEKERSKSAESPTSVTSETGSTPKKPLFGQAGHRRKTSRKPKE
DEVASEKKEQEKQEKVTEDEDAEYVASEKLTSTSTBOAHQOEAPESHAERELSAEYK
VELPSEVQSGSQSPSEKPAPLATEVPDEKIEVHDEVAEVAHVSTVEETSEOKTE
VEETAGSVPAEIVEMDAPEQAPAEVLVQLKTCVSGEDPTQGLSDPEKVLSTP
PEGVAVSEVMISQERMKVQSGPLKULFTSTGLKLSGKQKGRGGDEBSGHTVQ
PADSPSOEBOEGSSASPEPEITCLKEGLAEVQDGEAGEGATSDGKREGVY
PMASFKMVTPPKRVRPSPSDKDELIDYKSAITLSTESTASMOEMKGSVPEPKP
EERKXVDTYSWALICVSSSKBARAGSSSDEGPKAMGSHQXADAKGKRTG
TDGILASQEHDPQSSSPQAGSPTEBGEVSTWESFKRLVTRKSKSKLBEKSED

```


STAGSVEHSTPDTEPKESWWS IKKFIPIGRKKRPGDKQBOAPVEDAGPTGANNEDD
SDVPAVVFVLESDDAVERKWEAQQAQSAEQPEQKAATEVSKELSESQVHMAAAVAD
GTRAAITIIERSPSWTISASVTEPLEQVEAEALITBEVLEVEREVIABEEPTVTPELPE
NREAGDTTVSSEALTEPTAETAETAPLGAEEGTAEASAEETEMVSAVSQLTSPD
TTEATPTQVEVGGVPDIIEQERRTQELQVAEAKVBSQLPGTGGPDVLPQVORA
EAERPEQAEAGLKKETDVLVKVDAQEAKTRPPTQGVGOTTPESPEKAPQVTESI
ESSELVTCQAEITLAKVKSQEMVEQAIIPDSVETPTDSETGTSFVAPDFADPTQK
DRIVEIHENEVASGTQSGTBAEAVPAQKRPAPSSVFOEETKEQKMSMDLEHT
KGEVSEVETVLIUSLSEGTQADQYADEKTKVPFFEGLEGSIDGITVSEKVEVAL
DKEGTBEACCKDDALELQSAKPPSPVERMVQVREKTEAPETHVNEBKLEHET
LTAAREEVKLGENTANILETGETLEPAGAHVLEEKSSSEKNEKDFAHPCGEDAVPTGPD
LVTAAREEVKLGENTANILETGETLEPAGAHVLEEKSSSEKNEKDFAHPCGEDAVPTGPD
CONKSTPVIASATTKGLSDLEGEKTSILKWKSDREVQVACQEVKVSVAIEEDLEP
ENGILETSKSLQVNIITQAVDQFVTEETATEMLTSELQTOAHVIRKADSODAGQE
TEKEBEPQASQVETPTISAKESSESTAVGAHSDISKDSEASEKMTWVEGSETV
NQQLVEVLPSEEGGAGTKSPDEDDGHALLAERI BKSILVEPEDEKGDVDDPEN
ONSALADTDASGGLTKESPDTNGPKOKEKEDAQVEVLQBGKVHSESDKAITPOAEEL
OKQERESAKSELTES"

BASE COUNT 2051 a 1310 c 1708 g 1228 t

ORIGIN

Alignment Scores:
Pred. No.: 5,17e-143 Length: 6297
Score: 4231.50 Matches: 943
Percent Similarity: 69.07% Conservative: 187
Best local Similarity: 57.64% Mismatches: 398
Query Match: 52.42% Indels: 108
DB: 9 Gaps: 31

US-09-902-432-4 (1-1596) x AF001504 (1-6297)

QY 41 AspProAlaAspAla-----AspProAlaThrLysLeuProGlnLysAenGlyGlnLeu 58
Db 1 GACCCCGCCATCGCTGCTCGGACCCGCCACCAAGCTCTTACAGAGAATGGTTCAGCTG 60

QY 59 SerSerValAenGlyValAlaGluGlnGlyAspValHisValGlnGluGluAenGlnGlu 78
Db 61 TCACCATCAATGCGTAGCTGAGCAAGATGAGCTGACGCTCCAGGAGGGTGACCTAAAT 120

QY 79 GlyGln-----GluGluGluVal 84
Db 121 GGCCAGAAAGAGGCCCTGAACGGTCAAGGACCCCTAAACAGCCAGGAGGAAGAAGTC 180

QY 85 ValAspGluAspValGlyGlnArgGluSerGluAspValArgGluLysAspArgValGlu 104
Db 181 ATTGTTCACAGAGGTGGACAGAGAGACTCTCAAGATGTGAGCGAAAGAGACTCCGATAAA 240

QY 105 GluMetAlaAenSerThrAlaValGluAspIleThrLysAspGlyGlnGluGluThr 124
Db 241 GAGATGGCTACTAAGTCAGCGGTGTTCACGACATCACATGATGGCGAGGAGGACAC 300

QY 125 SerGluIleIleGluGlnIleProAlaSerGluAenValGluGluMetValGlnPro 144
Db 301 CCCGAAATATCGACAGATTCCTCTTCAGAAAGCAATTTAGAGAGCTTAACACACCC 360

QY 145 AlaGluSerGlnAlaAenAspValGlyPheLysLysValPheLysPheValGlyPheLys 164
Db 361 ACTGAGTCCCAGGCTAATGATATGGATTTTAAAGAGGTGTTTAAAGTTTGTGGCTTAAA 420

QY 165 PheThrValLysLysAspLysAenGluLysSerAspThrValGlnLeuLeuThrValLys 184
Db 421 TTCACGTGTGAAAAGGATAAGACAGAGAACCTGACACTCTCCAGCTACTCTCTGTGAAG 480

QY 185 LysAspGluGlyGluGlyAlaGluAlaSerValGlyAlaGlyAspHisGlnGluProSer 204
Db 481 AAAGATGAAGGGAGGAGCA-----GCAGGGGCTGGACACACAGGACCCGAGC 531

QY 205 ValGluThrAlaValGlyGluSerAlaSerLysGluSerGluLysGlnSerThrGlu 224
Db 532 CTT-----GGGGCTGGAGAGGAGCATCCAAAGAAAGCAACCCAAACAATCTACAGAG 585

QY 225 LysGlnGluGlyThrLeuLysGlnGlnSerSerThrGluLeuProLeuGlnAlaGlu 244
Db 1636 CAGGATGGGGAGGCTGAAGAAGGAGCTACTTCCGATGGAGAGAAAAAGAGAGAGGTGTC 1695

Db 586 AAACCCGAAGAGACCCCTGAAGCGTAGCAAAAGCCAGCAAAATTTCTCCCCAGCGGAA 645

QY 245 SerAspGlnAlaAlaGluGluAlaLysAspGluGlyGluGlnLysGlnGluLysGlu 264
Db 646 TCTGGCAAGCAGTG--GAGGAATCCAAAGAGGAAGAGAGAGAAACAAGAAAAAGAA 702

QY 265 ProThrLysSerProGluSerProSerProValAsnSerGluThrThrSerSerPhe 284
Db 703 CCTAGCAAGTCTCAGAAATCTCCGATAGTCCCGTACCAGTCAGAAACAGGATCAACCTTC 762

QY 285 LysLysPhePheThrHisGlyTTPAlaGlyTTPAlaGlyLysThrSerPheLysLysSer 304
Db 763 AAAAAATTTCTTACTCAAGTGGCGCGCTGGCGCAAAAGAACCACTTTTCAGAGACCG 822

QY 305 LysGluAspAspLeuGluThrAlaGluLysArgLysGluGlnGluAlaGluLysValAsp 324
Db 823 AAGGAGGATGAAGTGGAGGCTTCAGAGAGAGAAAAAGCAACAGAGCCAGAAAAAGTAGAC 882

QY 325 GluGluGluLysGluLysThrGluProAlaSerGluGlu----- 337
Db 883 ACAGAAAGAGACGGAAGGCGAGAGTTGGCTTCGAGAAAACCTGACCGCTCCGAGCAAGCC 942

QY 338 -----GlnGluProAlaGluAspThrAspGlnAlaArgLeuSerAlaAspTyrGluLys 355
Db 943 CACCCACAGGAGCGCGAGAAAGTGGCCAGAGCCCGGTTATCAGCTGAATATGAGAAA 1002

QY 356 ValGluLeuProLeuGluAspGlnValGlyAspLeuGluAlaSerSerGluGluLysCys 375
Db 1003 GTTGAGCTGCCCTCAGAGGAGCAAGTCAGTGGCTCGCAGGAGACCTTCTGAAGAGAAACCT 1062

QY 376 AlaProLeuAlaThrGluValPheAspGluLysMetGluAlaHisGln---GluValVal 394
Db 1063 GCTCCCGTTGGCGACAGAAGTGTGTGATGAGAAATAGAAAGTCCACCAAGAGAGGTGTG 1122

QY 395 AlaGluValHisValSerThrValGluLysThrGluGluGlnGlnGlyGlyGlyGlu 414
Db 1123 GCCGAAGTCCACGTACGACCGCTGGAGAGAGAAACCAAGAGAGCAG----- 1167

QY 415 AlaGluGlyValValValGluGlyThrGlyGluSerLeuProGluLysLeuAla 434
Db 1168 -----AAAACGGAGGTGGAGAAACACAGCAGGCTCTGTGCCAGCTGAAGAAATGGTT 1218

QY 435 GluProGlnGluValProGlnGluAlaGluProAlaGluGluLeuMetLysSerArgGlu 454
Db 1219 GAAATGGATGTCAGAACTCAGGAAGCGCAACCTGCCAAGAGAGCTGTGTAAGCTCAAAAGAA 1278

QY 455 MetCysValSerGlyGlyAspHisThrGlnLeuThrAspLeuSerProGluGluLysThr 474
Db 1279 ACGTGTGTTTCCGAGAGGAGCCCTACAGGGAGCTGACCTCAGTCTCTGATGAGAAGGTG 1338

QY 475 LeuProLysHisProGluGlyIleValSerGluValGluMetLeuSerSerGlnGluArg 494
Db 1339 CTGTCCAAACCCCGAGCGGTGTGTGAGTGTGGAATGCTGTCATCACAGGAGAGA 1398

QY 495 IleLysValGlnGlySerProLeuLysLysLeuPheSerSerSerGlyLeuLysLysLeu 514
Db 1399 ATGAAGTGTGAGGAGAGTCCACTAAAGAGCTTTTACCAGCACTGGCTTAAAAAGCTT 1458

QY 515 SerGlyLysLysGlnLysGlyLysArgGlyGlyGlyAspGluProGlyGluTyr 534
Db 1459 TCTGGAAGAAAAACAGAAAGGAAAAAG---GGAGAGGAGGACAGAGAAATCAGGGGAGCAC 1515

QY 535 GlnHisIleHisThrGluSerProGluSerAlaAspGluGlnLysGlyGluSerSerAla 554
Db 1516 ACTCAGGTTCACGCCATTTCTCCGACAGCCAGGAGAGCAAAAGGCGGAGAGCTCTGCC 1575

QY 555 SerSerProGluGluProGluGluThrThrCysLeuGluLysGlyProLeuGluAlaPro 574
Db 1576 TCATCCCTGAGGAGCCCGAGGAGATCAGTGTCTGGAAGAGGCTTAGCCGAGGTGACG 1635

QY 575 GlnAspGluAlaGluGluGlyThrThrSerAspGlyGluLysLysArgGluGlyIle 594
Db 1636 CAGGATGGGGAGGCTGAAGAAGGAGCTACTTCCGATGGAGAGAAAAAGAGAGAGGTGTC 1695

QY 595 ThrProTAlaSerPheLysLysMetValThrProLysArgValArgArgProSer 614
 Db 1696 ACTCCCTGGGCTTCATTCAAAAAGATGGTAGCCCAAGAACCGTTTAGAGCGGCTTCG 1755
 QY 615 GluSerAPlysGluGluLeuGluLysValLysSerAlaThrLysSerThrAsp 634
 Db 1756 GAAAGTGAATAAAGATAGTGGACAGGTCMAAGGGCTTACTTGCTTCACCGAG 1815
 QY 635 SerThrValSerGluMetGlnAspGluValLysThrValGluGluGluLysProGlu 654
 Db 1816 AGCAAGAGCTCTGAATATGCAGAAAGAAATGAAGGAGCGTGGAAGACCAAGCCGAA 1875
 QY 655 GluProLysArgArgValAspThrSerValSerTrpGluAlaLeuIleCysValGlySer 674
 Db 1876 GAACCAAAAGCCGAGGTGATACCTCAGATATCTTGGAAAGCTTTAATTTGTGGGATTA 1935
 QY 675 SerLysLysArgAlaArgLysAlaSerSerSerAspAspGluGluGlyProArgThrLeu 694
 Db 1936 TCCAAGAAAAGCAAGAGAGGTCCTCTCTGATGAGAAAGGGGACCAAAAGCAATG 1995
 QY 695 GlyGlyAspSerHisArgAlaGluAlaSerLysAspLysGluAlaGlyThrAspAla 714
 Db 1996 GGAGAGAGACCAACAGAAAGCTATAGAGCCGGAAGAAAGACAAAGAGACGAGCGG 2055
 QY 715 ValProAlaSerThrGlnGluGlnAspGlnAlaGlnGlySerSerSerProGluProAla 734
 Db 2056 ATCCTTGCTGTTCCCAAGAACATGATCCAGGCGAGGGAAGTTCTCCCGGAGCAAGCT 2115
 QY 735 GlySerProSerGluGluGluGluValSerThrTrpGluSerPheLysArgLeuValThr 754
 Db 2116 GGAACCCCAAGCAAGAGGAGGCGTTTCACCTGGAGTCATTAAAGTTAGTCAG 2175
 QY 755 ProArgLysLysSerLysSerLysLeuGluGluLysAlaGluAsp-----Ser 770
 Db 2176 CCAAGAAAAAATCAAAAGTCCAAAGCTGGAAGAAAGAAAGCAAGCTCCTAGCTGGGCT 2235
 QY 771 SerValGluGlnLeuSerThrGluIleGluProSerArgGluGluSerTrpValSerIle 790
 Db 2236 GGTTAGAAACATTCCACTCCAGACTGAAACCGGTAAAGAAATCTGGGCTCTCAATC 2295
 QY 791 LysLysPheIleProGluArgArgLysArgAlaAspGluLysGlnGlnAlaThr 810
 Db 2296 AAGAAGTTATTCCTGGAGCAAGAAAGGCAAGATGGGAAACAAAGCAAGCCCT 2355
 QY 811 ValGluAspSerGlyProValGluIleAsnGluAspAspProAlaValAlaVal 830
 Db 2356 GTTGAAGAGCGAGGCGCAAGGGGCCAAGAGATGACTCTGATGCTCCGGCGCGTGC 2415
 QY 831 ProLeuSerGluTrpAsnAlaValGluArgGluLysMetGlu-----AlaGlnGly 847
 Db 2416 CCTGCTGTAATGATCTGTAGAAAGGAGAAATGAGAGCACAGCAAGCCCAAAA 2475
 QY 848 AsnThrGluLeuProGlnLeuLeuGluAlaValTrpValSerGluGluLeuSerLysThr 867
 Db 2476 ACGCAGACAGCCCGCAAGAAAGCAGCAGCTGAGGTGTCAGAGAGCTCAGCGAGAG 2535
 QY 868 LeuValHisThrValSerValAlaValIleAspGlyThrArgAlaValThrSerValGlu 887
 Db 2536 CAGGTTCAATATGATGCAAGACTGTCGTCGAGCGGACAGGCGCACTACATATGAA 2595
 QY 888 GluArgSerProSerTrpLysSerAlaSerValThrGluProLeuGluLysIleThrAlaGly 907
 Db 2596 GAAAGTCTCTTCTTGATATCTGCTTCAGTACAGAACTCTTGAAACAAGTAGAGCT 2655
 QY 908 GluAlaMetProProValGluGluValThrGluLysAspIleIleIleIle---GluGluThr 926
 Db 2656 GAAGCGGCACTGTAACTAGAGGATTTGAAAGAAAGTAATTCACAAAGAAAGAACCC 2715
 QY 927 ProValLeuThrGlnThrLeuProGluGluLysAspAlaHisAspMetValThrSer 946
 Db 2716 CCAAGGTTACTGAACCTTGGCCAGAAACAGAGAGCCCGGCGCAACAGGTCGTTAGT 2775

QY 947 GluValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeuArgThr 966
 Db 2776 GAGCGGAATTGACCCCGAAGCTGTGACAGCTGCAGAAACTGACAGGCGCATGGGTGCC 2835
 QY 967 GluGluValThrGluAlaSerGlyValaGluGluThrThrAspMetValSerAlaValSer 986
 Db 2836 GAAGAGGAAACCGAAGCATCTGCTGTGAAGACACAGAAATGTGTCTCAGCACTGCC 2895
 QY 987 GlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGluValGluSer 1006
 Db 2896 CAGTTAACCGACTCCCGACAGACCAAGAGAGGCACTCCGCTGCAGAGAGGTGGAAAGT 2955
 QY 1007 GlyValLeuAspThrGluGluGluGluArgGlnThrGlnAlaIleLeuGluAlaValAla 1026
 Db 2956 GAGTAACTGACATGAAGAGACAGAGAGGCGAGCTCAAGAGTCTCCAGGCAAGTGCA 3015
 QY 1027 AspLysValLysGluGluGluSerGlnValProAlaThr-----Gln 1039
 Db 3016 GAAAGAGTGAAGAGGATCCAGCTGCTGGCACCGGTGGGCCAAGATGTCTCTCAG 3075
 QY 1040 ThrValGlnArgThrGlySerLysAlaLeuGluLysValGluGluValGluGluAspSer 1059
 Db 3076 CCTGTGCAGAGCA-----GAGCGCAAGAACCAAGACAGCAGGCT 3117
 QY 1060 GluValLeuAlaSerGluLysGluLysAspValMetProLysGlyProValGlnGluAla 1079
 Db 3118 GAAGGTCGGCTGTGAAGAAAGACGAGATGATGTTGAAGTAGATGCTCAGAGAGCA 3177
 QY 1080 GlyAlaGlnHisLeuAlaGlnGlySerGluThrGlyGlnAlaThrProGluSerLeuGlu 1099
 Db 3178 AAAAGTACCTTTACCAAGAGAGAGGTGGGCGAGACCAACCCAGAAAGCTTTGAA 3237
 QY 1100 ---ValProGluValThrAlaAspValAspHis-----ValAlaThrCysGlnVal 1115
 Db 3238 AAAGCTCTCAAGTACAGAGAGCATAGATCCAGTGAAGCTTTGTAACCATTTGCAACC 3297
 QY 1116 -----IleLysLeuGlnGlnLeu---MetGluGlnAlaValAlaProGlu 1129
 Db 3298 GAACCTTAGCTGGGGTAAATCACAGAGAGATGTGAATGGAACAGCATATCCCTTAGC 3357
 QY 1130 SerSerGluThrLeuThrAspSerGluThrAsnGlySerThrProLeuAlaAspSerAsp 1149
 Db 3358 TCGGTGAAACCCCTACAGACAGTGAAGTGAAGCAACCCCGTATGCCACTTTGAC 3417
 QY 1150 ThrAlaAspGlyThrGlnGlnAspGluThrIleAspSerGlnAspSerLysAlaThrAla 1169
 Db 3418 GCACAGCACAACCCAGAAAGACAGATTTGGAAATCCATAGAGAAATGAGGTCCGA 3477
 QY 1170 AlaValArgGlnSerGlnValThrGluGluGluAlaIleAlaThrAlaGlnGluGluPro 1189
 Db 3478 TCTGTACCCAGTCAAGGGGCAAGAACAGAGGCAAGTCTTCGACAAAGAAAGAGCCT 3537
 QY 1190 SerThrLeuProAsnAsnValProAlaGlnGluGluHisGlyGluGluProGly---Arg 1208
 Db 3538 CCAGCA-----CTTCCAGTTTGTGTTCCAGAGAGAACTAAAGAACATCAAGATGGA 3594
 QY 1209 AspValLeuGluProThrGlnGlnGluLeuThrAlaAlaValaValaProValLeuAlaLys 1228
 Db 3595 GACACTCTAGACATACAGATAAAGATGTCAGTGGAAGATGTATCCATCTGTCAAG 3654
 QY 1228 ThrGluValGlyGlnGluGluValaAspTrpLeuAspGlyGluLysValLys----- 1246
 Db 3655 ACTGAGGGGACTCAA-----GAGGCTACCAATGCTGATGAGAAACCAAGACGTA 3708
 QY 1247 -----GluGluGlnGluValaPheValHisSerGly-----ProAsnSerGlnLys 1261
 Db 3709 CCATTTTGAAGGACTTGAGGGGTCTATGACACAGCATTAAGCTAGTGGGAAG 3768
 QY 1262 AlaAlaAspValThrTyraAspSerGluValMetGlyValaGlyCysGlnGluLysGlu 1281
 Db 3769 GTCACTGAAGTTGCCCTTAAGAGTGAAGGACAGAAAGCTGAATGTAAAGAGATGAT 3828
 QY 1282 SerThrGluValGlnSer-----LeuSerLeuGluGluGluGluMetGlu 1296

Pred. No.: 4,236-137 Length: 104939
 Score: 4097.00 Matches: 917
 Percent Similarity: 68.28% Conservative: 181
 Best Local Similarity: 57.03% Mismatches: 391
 Query Match: 50.75% Indels: 119
 DB: 9 Gaps: 32

US-09-902-432-4 (1-1596) x AL590413 (1-104939)

QY 77 GlnGluGlyGlnGluGluValValAspGluAsp----- 88
 Db 9313 CAGCCGGGGCAAGAGTAGAGGCCCTGTGCGAGAAACAATTTCTTGATTTGTAATCA 9372
 QY 89 -----ValGlyGlnArgGluSer 94
 Db 9373 CCTTTTCTCTCTCCACCCGCCCTTTTGTATATAGTTGACAGAGAGACTCT 9432
 QY 95 GluAspValArgGluLysAspArgValGluGluMetAlaAlaSerThrAlaValGlu 114
 Db 9433 GAGAGTGTAGCAAAAGAGACTCCGATAAAGATGGCTAATGATCAGCGGTGTTCAC 9492
 QY 115 AspIleThrLysAspGlyGlnGluGluThrSerGluIleIleGluGlnIleProAlaSer 134
 Db 9493 GACATACAGATATATGGCAGAGAGAGACACCCGAAATTAATCGAACAAGATTCCTTCTCA 9552
 QY 135 GluAsnAsnValGluGluMetValGlnProAlaGluSerGlnAlaAsnAspValGlyPhe 154
 Db 9553 GAAAGCAATTTAGAAAGACTAACACCAACCACCTGAGTCCAGGCTATATATGATTTT 9612
 QY 155 LysLysValPheLysPheValGlyPheLysPheThrValLysLysAspLysAsnGluLys 174
 Db 9613 AAGAAAGGTGTTAAAGTTGTGGCTTAAATTCAGTGAAGAAAGATTAAGACAGAGAAAG 9672
 QY 175 SerAspThrValGlnLeuLeuThrValLysLysAspGluGlyGluGluAlaLys 194
 Db 9673 CCTGACACTGTCCAGCTACTCACTGAGAAAGATGAAGGGAGGAGCA----- 9723
 QY 195 ValGlyAlaGlyAspHisGlnGluProSerValGluThrAlaValGlyGluSerAlaSer 214
 Db 9724 GCGAGGGCTGGCGACCAACAGAGACCAGCCTT-----GGGGCTGAGAGAGAGCATCC 9777
 QY 215 LysGluSerGluLeuLysGlnSerThrGluLysGlnGluGlyThrLeuLysGlnGluGln 234
 Db 9778 AAGAAAGCGAACCCAAACATCTACAGAGAAACCGAAAGAGACCTGAAAGCGTGAAGCAA 9837
 QY 235 SerSerThrGluLeuProLeuGlnAlaGluSerAspGlnAlaIleGluGlnGluAlaLys 254
 Db 9838 AGCCACGCGAATAATTTCTCCCGCCGGAATCTGCCCAAGCATG---GAGGAATGCATA 9894
 QY 255 AspGluGlyGluGluLysGlnGluLysGluProThrLysSerProGluSerProSerSer 274
 Db 9895 GAGGAGGAGAAAGAGAAACAAAGAAAGAACTAGCAAGCTCAGCAATCTCCGACTAGT 9954
 QY 275 ProValAsnSerGluThrThrSerSerPheLysLysPhePheThrHisGlyTTPAlaGly 294
 Db 9955 CCGGTACCAAGTGAACAGAGATCAACCTTCAAAAAATTTCTCACTCAAGGTTGGCCGGC 10014
 QY 295 TrpArgLysLysThrSerPheLysLysSerLysGluAspSerLeuThrAlaGluLys 314
 Db 10015 TGGCGCAAAAGACCAAGTTTCAGGAAGCCGAAGAGATTAAGTGAAGCTTCAGAGAAAG 10074
 QY 315 ArgLysGluGlnGluAlaGluLysValAspGluGluGluLysGluLysThrGluProAla 334
 Db 10075 AAAAAGAAACAAAGCCAGAAAGTAAGACAGAAAGAAAGCGAAAGCGAGGTTGCC 10134
 QY 335 SerGluGlu-----GlnGluProAlaGluLysAsp 345
 Db 10135 TTCGAGAAACCTGACCGCTCCGAGCAAGCCACCAAGAGAGCGCGAGAAAGTCCAC 10194
 QY 346 GlnAlaArgLeuSerAlaAspTyrGluLysValGluLeuProLeuGlnGluAspGlnValGly 365
 Db 10195 GAGCCCGGTTATCAGCTGAATATGAGAAAGTTGAGCTGCCCTCAGAGAGCAAGTCACT 10254

QY 366 AspleuGluAlaSerSerGluGluLysCysAlaProLeuAlaThrGluValPheAspGlu 385
 Db 10255 GAGTCGAGGAGACCTTCTGAAGAGAAACCTCTCCGTGGGAGCAAGATTTGATGAG 10314
 QY 386 LysMetGluAlaHisGln---GluValValAlaGluValHisValSerThrValGluLys 404
 Db 10315 AAAATGAAAGTCCACCAAGAAAGAGTTGTGGCCGGAAGTCCAGTCAAGACCGTGGAGAG 10374
 QY 405 ThrGluGluGlnGlnGlyGlyGlyGluAlaGluGlyValValGluGlyThr 424
 Db 10375 AGAACCGAAGGACAG-----AAACGAGAGTGAAGAAACA 10410
 QY 425 GlyLysLeuProProGluLysLeuAlaGluProGlnGluValProGlnGluAlaGlu 444
 Db 10411 GCGAGGTCTGTGCGCAAGTGAATGGTTGAATGATGCAAACTCAGAGAACTGAA 10470
 QY 445 ProAlaGluGluLeuMetLysSerArgGluMetCysValSerGlyGlyAspHisThrGln 464
 Db 10471 CCTGCCAAGAGCTGTGTGAACCTCAAGAAACGTGTGTTCCGAGAGAGACCTTACACAG 10530
 QY 465 LeuThrAspLeuSerProGluGluLysThrLeuProLysHisProGluGlyIleValSer 484
 Db 10531 GAGCTGACCTCAGCTCAGATGAGAGAGTGTCTCCAAACCCCGAAGCGGTGTGAGT 10590
 QY 485 GluValGluMetLeuSerSerGlnGluArgIleLysValGlnLysSerProLeuLysLys 504
 Db 10591 GAGGTGAATATGCTGTCACTACAGAGAGAAATGAAGGTGCAGAGGAAGTCACTAAAGAAAG 10650
 QY 505 LeuPheSerSerSerGlyLeuLysLysLeuSerGlyLysLysGlnLysGlyLysAspGly 524
 Db 10651 CTTTTCACGACAGCTGGCTTAAAAAGCTTCTGGAAGAAACGAAAGGAGAAAGA--- 10707
 QY 525 GlyIleGlyAspGluGluProGluGlyGlyGlyGlnHisIleHisThrGluSerProGluSer 544
 Db 10708 GAGAGAGAGAGAGAGAGAAATCAAGAGAGACACTCAGTGTCCAGCCGATTTCCGAGACAG 10767
 QY 545 AlaAspGluGlnLysGlyGluSerSerAlaSerSerProGluGluProGluGluThrThr 564
 Db 10768 CAGAGAGAGCAAAAGAGAGAGAGCTGTGCTCATCTCCCTGAGAGACCCGAGAGAGTCAAG 10827
 QY 565 CysLeuGluLysGlyProLeuGlnGluAlaProGlnAspGlyGluAlaGluGluGlyThrThr 584
 Db 10828 TGTCTGAAAAGGCTTTCAGAGTCCAGAGTCCAGAGAGTGGGAAGCTGAAAGAGACTACT 10887
 QY 585 SerAspGlyGlyLysLysArgGlyGlyIleThrProTTPAlaSerPheLysLysMetVal 604
 Db 10888 TCCGATGAGAGAAAGAAAGAAAGAGTGTCACTCCCTGGGCTATCTCAAAAAGATGGTG 10947
 QY 605 ThrProLysLysArgValArgArgProSerGluSerAspLysGluGluGluLeuGluLys 624
 Db 10948 ACGCCCAAGAACGTTTATGAGAGCGCTTCGAAAGTGAATGAAGAGATGAGTGAACAAG 11007
 QY 625 ValLysSerAlaThrLeuSerSerThrAspSerThrValSerGluMetGlnAspGluVal 644
 Db 11008 GTCAAGAGGCTTACTTGTCTTCCACCGAGACACAGCTCTGAAATTCAGAAAGAAATG 11067
 QY 645 LysThrValGlyGluGlnGluLysAspProGluGluProLysArgArgValAspThrSerVal 664
 Db 11068 AAAGGAGCGGTGAAGAGAGCAAAAGCCGGAAGAAACCAAGCCGCAAGTGGATCTCAGTA 11127
 QY 665 SerTPGluAlaLeuIleLysValGlySerSerLysArgAlaArgLysAlaSerSer 684
 Db 11128 TCTTGGGAAGCTTTAATTTGTGTGAGATCTCAAGAAAGAGCAAGAGAGAGGTCCTT 11187
 QY 685 SerAspAspGluGlyGlyProArgThrLeuGlyLysSerHisArgAlaGluGluAla 704
 Db 11188 TCTGATAGAGAGAGGAGACCAAAAGCAATGGAGAGAGACCAACAAAGCTGATAGAGCC 11247
 QY 705 SerLysAspLysGluAlaGlyThrAspAlaValProAlaSerThrGlnGluGlnAspGln 724
 Db 11248 GGAAAAAGACAAAGAGACGAGGAGACAGCGGAGCTTGTGCTCCAGAAACATGATCA 11307
 QY 725 AlaGlnGlySerSerSerProGluProAlaGlySerProSerGluGlyGluGlyAlaSer 744

DB	11308	GGGCAAGGGAAGTTCCTCCCGGAGCAAGCTGGAAAGCCCTACCGAAGGGAGGGCGGTTCCTC	11367
QY	745	ThrTrpGluSerPheLysArgLeuValThrProArgLysLysSerLysSerLysLysLeuGlu	764
DB	11368	ACCTGGGAGTCATTTAAAGGTGTAGTCAGCCAGAAAATAATCAAGTCCAGAGTCGAA	11427
QY	765	GlutysAlaGluAsp-----SerSerValGluGlnLeuSerThrGluIleGlu	780
DB	11428	GAGAAAAGCGAAGACTCCATAGCTGGGTCTGGTGTAAGAACATTCCACTCCAGACACTGAA	11487
QY	781	ProSerArgGluGluSerTrpValSerIleLysLysPheIleProGlyArgArgLysLys	800
DB	11488	CCCGTAAAGAAAGAAATCTCGGTCTCAATCAAGAAAGTTATTCTCTGGACGAAGGAAGAA	11547
QY	801	ArgAlaAspGlyLysGlnGluGlnAlaThrValGluAspSerGlyProValGluIleAsn	820
DB	11548	AGGCAGATGGGAAACAGACACGACCCCTGTTGAAAGACGACGAGGCCAACAGGGGCCAAC	11607
QY	821	GluAspAspProAsnValProAlaValValProLeuSerGluTyrAsnAlaValGluArg	840
DB	11608	GAAGATGACTCTGATGTCCTCGGCCGTGCTCTGTCTGAGTATGATGCTGTAGAAAGG	11667
QY	841	GlutysMetGlu-----AlaGlnGlyAsnThrGluLeuProGlnLeuLeuGlyAla	857
DB	11668	GAGAAAATGGAGGCACAGCAAGCCCAAAAAAGCGACAGACGCCGACGACGAAGGCGACGC	11727
QY	858	ValTyrValSerGluGluLeuSerLysThrLeuValHisThrValSerValAlaValIle	877
DB	11728	ACTGAGGTGTCACAGGAGCTCAGCGAGAGTCAGGTTCATATGATGGCAGCAGCTGTGCT	11787
QY	878	AspGlyThrArgAlaValThrSerValGluGluArgSerProSerTrpIleSerAlaSer	897
DB	11788	GACGGCACGAGGCGAGCTACCATTTATGAAGAAAGTCTCTCTCTTGGATATCTGCTTCA	11847
QY	898	ValThrGluProLeuGluHisThrAlaGlyGluAlaMetProProValGluGluValThr	917
DB	11848	GTGACAGAACCTCTTGAACAAGTAGTAAGACTGAAAGCGCAGCTGTTAACTGAGGAGGTATTG	11907
QY	918	GluLysAspIleAla--GluGluThrProValLeuThrGlnThrLeuProGluGly	936
DB	11908	GAAAGAGAAGTAATTGCAGAGAAGAACCCGCCGTTACTGAACTCTGCCAGAGAAC	11967
QY	937	LysAspAlaHisAspAspMetValThrSerGluValAspPheThrSerGluAlaValThr	956
DB	11968	AGAGAGGCCCGGGGCGACACAGCTCGTTAGTGAGGGCGGAATTGACCCCGAAGCTGTGACA	12027
QY	957	AlaThrGluThrSerGluAlaLeuArgThrGluGluValThrGluAlaSerGlyAlaGlu	976
DB	12028	GCTCGAAGAACTGCAGGCGCAATGGGTGCCGAAGAAGAACCGAAGCATCTGCTGCTGAA	12087
QY	977	GluThrThrAspMetValSerAlaValSerGlnLeuThrAspSerProAspThrThrGlu	996
DB	12088	GAGACCACAGAAATGTTGTCTCAGCATCTCCCAAGTTAACCGACTCCCAAGACACACAGAG	12147
QY	997	GluAlaThrProValGlnGluValGluSerGlyValLeuAspThrGluGluGluGluArg	1016
DB	12148	GAGGCCACTCCGTCGAGAGGTGGGAAGGTGGCGTACCTGACATAGAAGACCAAGAGAGG	12207
QY	1017	GlnThrGlnAlaIleLeuGlnAlaValAlaAspLysValLysGluLeuSerGlnValPro	1036
DB	12208	CGGACTCAAGAGGTCTCTCCAGCAGTGCAGAAAAGTGAAGAGGAATCCAGCTGCTCT	12267
QY	1037	AlaThr-----GlnThrValGlnArgThrGlySerLysAlaLeu	1049
DB	12268	GGCACCGTGGCCAGAGATGTGCTCAGCCTGTGCAGAGACA-----	12312
QY	1050	GlutysValGluGluValGluGluAspSerGluValLeuAlaSerGluLysGluLysAsp	1069
DB	12313	--GAGGCAGAAAGACCAAGAGAGCAGCTGAAGCGTCGGGTCTCAAGAAAGACCGAT	12369
QY	1070	ValMetProLysGlyProValIcGlnAlaGlyAlaGluHisLeuAlaGlnGlySerGlu	1089

Db	12370	GTAGTGTGTTGAAAGTAGTAGTCTCAGGAGGCCAAAACTGAGCCCTTTTACAAAGGGAAGGTG	12422
Qy	1090	ThrGlyGlnAlaThrProGluSerLeuGlu--ValProGluValThrAlaAspValAsp	1108
Db	12430	GTGGGGCAGACCCACCCAGAAAGCTTTGAAAAGCTCTCAAGTCCACAGAGCATAGAG	12489
Qy	1109	His-----ValAlaThrCysGlnVal-----IleLysLeuGlnGln	1120
Db	12490	TCCAGTGTAGCTTGTAACCACTTGTCAAGCCGAAACCTTAGCTGGGGTAAATAACAGGAG	12549
Qy	1121	Leu---MetGluGlnAlaValAlaProGluSerSerGluThrLeuThrAspSerGluThr	1139
Db	12550	ATGGTGATGGAAACAGGCTATCCCCCTGACTCGGTGGAAACCCCTACAGACAGTGAGACT	12609
Qy	1140	AsnGlySerThrProLeuAlaAspSerAspThrAlaAspGlyThrGlnGlnAspGluThr	1159
Db	12610	GATGGAGACACCCCGTAGCCGACTTTGACGCACCCAGGCACAAACCCAGAAAGACGAGATT	12669
Qy	1160	IleAspSerGlnAspSerLysAlaThrAlaAlaValArgGlnSerGlnValThrGluGlu	1179
Db	12670	GTGGAAATCCATGAGGAGCAATGAGTGCATCTGGTACCCAGTCAGGGGGCACAGAAGCA	12729
Qy	1180	GluAlaAlaThrAlaGlnLysGluGluProSerThrLeuProAsnAsnValProAlaGln	1199
Db	12730	GAGGCAGTTCCTGCACAGAAAGAGAGGCGCTCCAGCA---CCTTCAGTTTTGTGTCCAG	12786
Qy	1200	GluGluHisGlyGluGluProGly---ArgAspValLeuGluProThrGlnGlnGluLeu	1218
Db	12787	GAAGAAACTAAGAACAATCAAGATGGAAGACACTCTAGAGCATACAGATAAGAGGTG	12846
Qy	1219	ThrAlaAlaValProValLeuAlaLysThrGluValGlyGlnGluGlyGluValAsp	1238
Db	12847	TCAGTGGAAACTGTATCCATTCTGTCAAGACTGAGGGGACTCAA-----GAGGCTCAC	12900
Qy	1239	TrpLeuAspGlyGluLysValLys-----GluGluGlnGluValIlePheVal	1253
Db	12901	CAGTATGCTGATGAGAAACCAAGACAGTACCATTTTTCGAAGGACTTGAGGGGTCTATA	12960
Qy	1254	HisSerGly-----ProAsnSerGlnLysAlaAlaAspValThrTyrAspSerGluVal	1271
Db	12961	GACACAGGCATACAGTCAAGTCGGGAAAGGTCACTGAAAGTTGCCCTTAAAGGTGAAGGG	13020
Qy	1272	MetGlyValAlaGlyCysGlnLysGluSerThrGluValGlnSer-----	1287
Db	13021	ACAGAAGAAGCTGAATGTAAAAAGGATGATGCTCTTGAACTCGACAGTCACGCTAAGTCT	13080
Qy	1288	---LeuSerLeuGluGluGlyGluMetGluThrAspValGluLysGluLysArgGluThr	1306
Db	13081	CCTCCATCCCCGTGGAGAGAGATGTTAGTTCAAGTCGAAGGGGAGAAACAGAAGCA	13140
Qy	1307	LysProGluGlnValSerGluGluGly---GluGlnGluThrAlaAlaProGluHisGlu	1325
Db	13141	GAGCCACCCCATGTGAATGAAGAGAGAGCTTGAGCACGAAACAGCTGTTACCGTATCTGAA	13200
Qy	1326	GlyThrTyrGlyLysProValLeuThrLeuAspMetProSerSerGluArgGlyLysAla	1345
Db	13201	GAGGTCAGTAAGCAGCTCCTCCACAGACAGTGAATGTGCCCATCATAGTGGGGCAAGGAA	13260
Qy	1346	LeuGlySerLeuGlyGlySer---ProSerLeuProAspGlnAspLysAlaGlyCysIle	1364
Db	13261	GTCAGCATTTTGGGAAGGAAGGCCCTCCTCCCTGCGCTCAAGGAGGCGCATGTATGCACC	13320
Qy	1365	GluValGlnValGlnSerLeuAspThrThrValThrGlnThrAlaGluAlaValGluLys	1384
Db	13321	AAATTCAGTTTCAGAGCTCTGAGGCATCATTCATCTTAACAGCGGCTGCAGAGAGGAA	13380
Qy	1385	ValIle-----GluThrValValIleSerGluThrGlyGluSerProGluCysValGly	1402
Db	13381	AGGTCTTTAGGAAAACTGCCAACATTTGAACAGGTGAACAGCTTCGAGCCTCGAGGT	13440
Qy	1403	AlaHisLeuLeuProAlaGluLysSerSerAlaThrGlyGlyHisTrpThrLeuGlnHis	1422
Db	13441	GCACATTTTAGTTCTGGGAAGAGAAATCTCTCGAAAAAATGAAGACTTTGCCGCTCATCCA	13500

Qy 1423 AlagiuaSpThrValProleuGlyProGluSerGlnAlaGluSerIleProIleIleVal 1442
 Db 13501 GGGGAAGATGCTGTGCCCCACAGGGCCGACTGTCCAGGCAAAATCCAGACAGTGTAGTA 13560
 Qy 1443 ThrProAlaProGluSerThrIleuHisProAspIleuGlnGlyIleSerIleAspGln 1462
 Db 13561 TCTGCTACTACCAAGAAAGGCTTAAGTTCGACCTGGAAAGAGAGAAACCACTACTG 13620
 Qy 1463 ArgGluArgSerGluGluGluIleuAspIleuGlnProAspAlaGlyProAspIleuGln 1482
 Db 13621 AAGTGAATCTCAGTAACTCATTCATGACGAGTTGCTTGCACAGAGGTC-----AAAGT 13674
 Qy 1483 SerThrAlaIleGluIleValIleuIleuAspIleuGlnProGlu-----IleleuGluIleuGln 1500
 Db 13675 AGCTGAGCAATGAGGATTTA-----GAGCCTGAAATGCGATTTTGGCACTTGAG 13725
 Qy 1501 SerIleuSerAspIleuIleValIleuAspIleuGlnIleGlnThrAlaValAspIleuIleuAsp 1520
 Db 13726 ACCAAAGAGATTAAGTCTGTCCAAACATCATCCAGACAGCGTTGACCAAGTTGTACGT 13785
 Qy 1521 Thr---GluThrAlaProGluThrHisAlaIleuAspSerGlnThrGlnValProAlaCys 1539
 Db 13786 AAGAAAGAAACAGCCACCCAAATGTTGAGCTGCTGACTTACAGACACAACTCAGCGATA 13845
 Qy 1540 ArgIleuAspSerArgGluProAspArgCysTyrPThrIle-----MetIle 1554
 Db 13846 AAAGCTGACAGCCAGACGCTGGACAGAAAGAGAAAGAGAGAGAGAACTCAGGCC 13905
 Qy 1555 AspAla-IleuMetIleuHisProValProGlnProAlaGlu----- 1567
 Db 13906 TCTGCACAGATGAACACCAATTACTTACGCCAAAGAGAGAGTCAAGTCAACCGCAGTG 13965
 Qy 1568 -AspIleuGlnValIleuThrValIleuGlnIleuAlaIleuAlaIleuProArgCysIleuProArg 1587
 Db 13966 GAGCAAGCAGCATTTCTATTTTCCAAAGACATGACATGAAAGCCCTCAGAAAGAACCACTGACT 14025
 Qy 1587 GieugIleuIleuAspAlaProVal 1594
 Db 14026 GTTGAAGTAGAAGGTTCCACTG 14047

RESULT 15
 AC023201 171158 bp DNA linear HTG 24-AUG-2002
 LOCUS Homo sapiens chromosome 6 clone RP11-193D24 map 6, WORKING DRAFT
 DEFINITION AC023201
 ACCESSION AC023201.2 GI:7139764
 VERSION HTG; HTGS_PHASE1; HTGS_DRAFT.
 KEYWORDS human.
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 171158)
 Birren, B., Linton, L., Nusbaum, C. and Lander, E.
 Homo sapiens chromosome 6, clone RP11-193D24
 Unpublished
 2 (bases 1 to 171158)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
 Anderson, S., Baldwin, J., Barna, N., Beckerly, R., Beda, F.,
 Boguslavsky, L., Bouknight, B., Brown, A., Burkett, G., Castle, A.,
 Choepey, Y., Colangelo, M., Collins, S., Collamore, A., Cooke, P.,
 Chevalier, K., Dewar, K., Domingo, M., Doyle, M., Fenece, J.,
 Ferreira, P., Fitzhugh, W., Forrest, C., Gage, D., Galagan, J.,
 Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,
 Howland, J., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,
 Landers, J., Lehoczy, J., Levine, R., Liu, C., Liu, G., Locke, K.,
 Macdonald, P., Margulis, N., McEwan, P., McGurk, A., McKernan, K.,
 McPheters, R., Meldrum, J., Menus, L., Morrow, J., Naylor, J.,
 Norman, C. H., O'Connor, T., O'Donnell, P., Oliver, T. M., Peterson, K.,
 Pierre, N., Pisan, C., Pollara, V., Raymond, C., Riley, R., Rothman, D.,
 Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N.,
 Stojanovic, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,

TITL
 JOURNAL
 REFERENCE
 AUTHORS

TITL
 JOURNAL
 COMMENT

Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W. J.,
 Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (09-FEB-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 171158)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
 Anderson, S., Baldwin, J., Barna, N., Beckerly, R., Beda, F.,
 Boguslavsky, L., Bouknight, B., Brown, A., Burkett, G.,
 Campopiano, A., Castle, A., Choepey, Y., Colangelo, M., Collins, S.,
 Collamore, A., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J. S.,
 Dodge, S., Domingo, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,
 Galagan, J., Gardyna, S., Ginde, S., Coyette, M., Graham, L.,
 Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,
 Howland, J., Johnson, R., Jones, C., Kann, L., Karatas, A.,
 Klein, J., Lehoczy, J., Levine, R., Landers, J., Lehoczy, J.,
 Levine, R., Liu, G., Locke, K., Macdonald, P., Margulis, N.,
 McCarthy, M., McEwan, P., McGurk, A., McKernan, K., Menus, L.,
 Meldrum, J., Menus, L., Morrow, J., Miranda, C., Mlenga, V., Morrow, J.,
 Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,
 O'Neil, D., Oliver, T. M., Oliver, J., Peterson, K., Pierre, N.,
 Pisan, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
 Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
 Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
 Testaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,
 Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J.,
 Young, G., Zainoun, J., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Mar 1, 2000 this sequence version replaced gi:6957769.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WTHR
 Web site: http://www-seg.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L6150
 Center clone name: 193 D 24
 ----- Summary Statistics
 Sequencing vector: M13; M77815; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 159689 bases at least Q40
 Consensus quality: 164965 bases at least Q30
 Consensus quality: 167080 bases at least Q20
 Insert size: 176000; agarose-IP
 Insert size: 168658; sum-of-contigs
 Quality coverage: 4.8 in Q20 bases; agarose-IP
 Quality coverage: 5.0 in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 26 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 1187: contig of 1187 bp in length
 * 1188 1287: gap of 100 bp
 * 1288 3150: contig of 1863 bp in length
 * 3151 3250: gap of 100 bp
 * 3251 4993: contig of 1743 bp in length
 * 5000 5099: gap of 100 bp
 * 5100 6170: contig of 1071 bp in length
 * 6171 6270: gap of 100 bp
 * 6271 7887: contig of 1617 bp in length
 * 7888 7987: gap of 100 bp
 * 7988 9424: contig of 1437 bp in length

Qy	957	AlaThrGluThrSerGluAlaLeuArgThrGluGluValThrGluAlaSerGlyAlaGlu	976
Db	90016	GCTGCAGAAACTCTGAGGGCCCATTTGGTGCCGAAGAAGAAACCGAAGCATCTGCTGCTGAA	90075
Qy	977	GluThrThrAspMetValSerAlaValSerGlnLeuThrAspSerProAspThrThrGlu	996
Db	90076	GAGACCACAGAAATGGTGACAGCTCTCCAGATTAAACCGACTCTCCACAGACACACAGAG	90135
Qy	997	GluAlaThrProValGlnGluValGluSerGlyValLeuAspThrGluGluGluAlaArg	1016
Db	90136	GAGGCCACTCTCGGTGCAGAGGTGGAAGGTGGCGTACCTGCATAGAGAAGACAGAGAGG	90195
Qy	1017	GlnThrGlnAlaIleLeuGluAlaValAlaAspLysValLysGluGluSerGlnValPro	1036
Db	90196	CGGAICTCAAGAGGTCTCTCCAGCGAGTCGACAAAAGTGAAGAAGGAATCCAGCTGCCT	90255
Qy	1037	AlaThr-----GlnThrValGlnArgThrGlySerLysAlaLeu	1049
Db	90256	GGCACCGGTGGCCGACGAAGATGTGCTTCAGCTGTGCAGAGCA-----	90300
Qy	1050	GluLysValGluGluValGluAspSerGluValLeuAlaLaserGluLysGluLysAsp	1069
Db	90301	---GAGGCAGAAAGACACAGAGACAGCGCTGAAGCGTCGGTCTCAAGAAAGACCGAT	90357
Qy	1070	ValMetProLysGlyProValGlnGluAlaGlyAlaGluHisLeuAlaGlnGlySerGlu	1089
Db	90358	GTAGTGTGGAAGTAGATGCTCAGGAGCAAAACTAGCGCTTTTACAAAGGGAAGGTG	90417
Qy	1090	ThrGlyGlnAlaThrProGluSerLeuGlu---ValProGluValThrAlaAspValAsp	1108
Db	90418	GTGGGGCAGACACCCCAAGAAAGCTTTGAAAAGCTCTCAAGTCACAGAGACATAGAG	90477
Qy	1109	His-----ValAlaThrCysGlnVal-----IleLysLeuGlnGln	1120
Db	90478	TCCAGTGAGCTGTGAACCACTGTCAAGCCGAAACCTTAGCTGGGGTAAATATCACAGAG	90537
Qy	1121	Leu---MetGluGlnAlaValAlaProGluSerSerGluThrLeuThrAspSerGluThr	1139
Db	90538	ATGTGTATGGAACAGGCTATCCCCCTGACTCGGTGGAAACCCCTACAGACAGTGAGACT	90597
Qy	1140	AsnGlySerThrProLeuAlaAspSerAspThrAlaAspGlyThrGlnGlnAspGluThr	1159
Db	90598	GATGGAACACCCCGTAGCCGACTTTGACGCACCCAGCCACACCCCAAGAAAGACGAGATT	90657
Qy	1160	IleAspSerGlnAspSerLysAlaThrAlaValArgGlnSerGlnValThrGluGlu	1179
Db	90658	GTGGAAATCCATGAGGAGAATGAGGTGCATCTGGTACCCAGTCTAGGGGGCACAGAAGCA	90717
Qy	1180	GluAlaAlaThrAlaGlnLysGluGluProSerThrLeuProAsnAsnValProAlaGln	1199
Db	90718	GAGCAGTTCTGCACAGAAAGAGAGGCCCTCCAGCA---CCTCCAGTTTGTGTCTCCAG	90774
Qy	1200	GluGluHisGlyGluGluProGly---ArgAspValLeuGluProThrGlnGlnGluLeu	1218
Db	90775	GAAGAAACTAAAGAACCATCAAAAGTAGGAAGACACTCTAGAGCATACAGATAAAGAGGTG	90834
Qy	1219	ThrAlaAlaAlaValProValLeuAlaLysThrGluValGlyGlnGlnGlyValAsp	1238
Db	90835	TCAGTGGAACTGTATCCATTCTGTCAAAAGACTGAGGGGACTCAA-----GAGCGTGAC	90888
Qy	1239	TrpLeuAspGlyGluLysValLys-----GluGluGlnGluValIlePheVal	1253
Db	90889	CAGTATGCTGATGAGAAACCAAGACAGCTACCATTTTTCGAAGGACTTGAGGGGTCTATA	90948
Qy	1254	HisSerGly-----ProAsnSerGlnLysAlaAlaAspValThrTyAspSerGluVal	1271
Db	90949	GACACAGCATTAACAGCTCAGTCGGGAAAGGTCTCAAGTGTGCCCTTAAAGGTGAAGGG	91008
Qy	1272	MetGlyValAlaGlyCysGlnGluLysGluSerThrGluValGlnSer-----	1287
Db	91009	ACAGAAGNAGCTGAATGTAAAGAAAGATGATGCTCTTGAACCTGCAGAGTCACGCTAAGTCT	91068
Qy	1288	---LeuSerLeuGluGluGlyGluMetGluThrAspValGluLysGluLysArgGluThr	1306

Db	91069	CCTCATCCCGCTGGAGAGAGATG	GTAGTTCAAGTCTGAAAGGGAGAAACAGAACCA	91126
Qy	1307	LysProGluGlnValSerGluGluGly	---GluGlnGlnThrAlaAlaProGluHisGlu	1325
Db	91129	GAGCCAAACCCATGTGAATTAAGAGAGAAGCTT	CAGACACGAAACAGCTGTTTACCGTATCTGAA	91188
Qy	1326	GlyThrTyGlyLysProValLeuThrLeuAspMet	ProSerSerGluArgGlyLysAla	1345
Db	91189	GAGGTCAGTAGCAGCTCTCCAGACAGCTGAATGT	GCCCATCATAGATGGGGGCAAAAGAA	91248
Qy	1346	LeuGlySerLeuGlyGlySer---	ProSerLeuProAspGlnAspLysAlaGlyCysIle	1364
Db	91249	GTACAGCATTTGGGAAGGAAGCCCTCTCCCTGC	TAGTCAAGAGGAGCGCATGCAACC	91308
Qy	1365	GluValGlnValGlnSerLeuAspThrThrValThr	GlnThrAlaGluAlaValGluLys	1384
Db	91309	AAATTCAGATTTCAGAGCTCTGTAGCGCATCAT	TTCACTTAAACAGCGCGTGCAGAGGAGAA	91368
Qy	1385	ValIle-----GluThrValValIleSerGluThr	GlyGlyGluSerProGluCysValGly	1402
Db	91369	AAGGTCITTAGGAGAACTGCCAACATTTAGAAC	ACAGTGAACGTTGGAGCCTCCAGGT	91428
Qy	1403	AlaHisLeuLeuProAlaGluLysSerSerAlaThr	GlyGlyHisTrpThrLeuGlnHis	1422
Db	91429	GCACATTTAGTTCTGGAAGAGAAATCTCTCAAAA	AAATGAAGACTTTGCCGCTCATCCA	91488
Qy	1423	AlaGluAspThrValProLeuGlyProGluSerGln	AlaGluSerIleProIleIleVal	1442
Db	91489	GGGAAGATGCTGTGCCACAGGGCCGACCTGT	CAGGCAAAATGCACACCAAGTATAGTA	91548
Qy	1443	ThrProAlaProGluSerThrLeuHisProAspLeu	GlnGlnGlyGluIleSerAlaSerGln	1462
Db	91549	TCTGCTACTACCAAGAAAGGCTTAAGTCCGAC	CTCGAAGGAGAGAAACCATCACTG	91608
Qy	1463	ArgGluArgSerGluGluAspLysProAspAlaGly	ProAspAlaAspGlyLysGlu	1482
Db	91609	AAGTGGAAAGTCAGATGAAGTCGATGAGCAGGT	TGCTTCCAGAGGTC-----AAAGTG	91662
Qy	1483	SerThrAlaIleGluLysValLeuLysAlaGluPro	Glu-----IleLeuGluLeuGlu	1500
Db	91663	AGTGTAGCAATTGAGGATTTA-----GAGCT	GNAATGGGATTTTGGAACTTGAG	91713
Qy	1501	SerLysSerAsnLysIleValLeuAsnValIleGln	ThrAlaValAspGlnPheAlaArg	1520
Db	91714	ACCAAAAGCACTAACTTGTCACAAACATCATCC	AGACAGCGGTTGACCACTTTGTACGT	91773
Qy	1521	Thr---GluThrAlaProGluThrHisAlaIle	TrpSerGlnThrGlnValProAlaCys	1539
Db	91774	ACAGAAGAAAACAGCCACCGAAATGTGACGTCT	GATGTTGAGTTACAGACAAAGCTCACG	91833
Qy	1540	ArgLeuAspSerArgGluProAsnArgCysTrpThr	Lys-----MetLys	1554
Db	91834	AAAGCTGACCCAGGAGCGCTGGACGAAACCG	GAGAAAGAGGAGAGAACTCAGGCC	91893
Qy	1555	AspAla-LysMetLysHisProValProGlnPro	ArgGlu-----	1567
Db	91894	TCTGCACAGGATGAACACCAATTACTTTCAGCC	AAAGAGGAGTCAAGCGCAGTG	91953
Qy	1568	-AspLeuGlnValLeuThrValLeuGluAlaTrp	AlaGlnProArgLysCysLeuProAr	1587
Db	91954	GGACAGCACATTTCTGATATTTCCAAAGACAT	GATGAGTGAAGCCTCAGAAAGACCATGACT	92013
Qy	1587	GLeuGlnLeuLysAlaProVal	1594	
Db	92014	GTTCAGGTAGAAGGTTCCACTG	92035	

Search completed: December 13, 2002, 03:13:17
Job time : 9259 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

QM protein - nucleic search, using frame_plus_p2n model
Run on: December 13, 2002, 00:16:41 ; Search time 619 Seconds
(without alignments)
5806.445 Million cell updates/sec

Title: US-09-902-432-4
Perfect score: 8073
Sequence: 1 MAGSSTQRSPEQPGSDT.....AWAQRKCLPRLQKAPVSK 1596

Scoring table:
BLOSUM62 Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgn2.1/USPTO.spool/US0902432/runat.12122002.141451.25790/app.query.fasta_1.1735
-DB=N Geneseq 101002 -OPMT=fastap -SUFFIX=mg -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DLOCAL=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptp -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0902432 @CNG 1 1 488 @runat.12122002.141451.25790 -NCPU=6 -ICPU=3
-NO XLPXY -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
N Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	8041	99.6	5200	18 AAV02302	Rat tumour suppress
2	6001	74.3	5074	18 AAV02301	Rat tumour suppress
3	4340.5	53.8	6614	24 ABN59861	Novel human coding
4	4300.5	53.3	6608	24 ABN59842	Gene #3440 used to
5	4300.5	53.3	6608	24 ABL61938	Colon adenocarcino
6	4281	53.0	6605	19 AAV23545	Human gravin poly
7	4281	53.0	6605	21 AAV74903	Human gravin codin
8	4226.5	52.4	6886	23 AAS85205	DNA encoding novel
9	642	8.0	8307	23 ABL23161	Drosophila melanog
10	639.5	7.9	11597	23 ABL23160	Drosophila melanog
11	565.5	7.0	18506	23 ABL05592	Drosophila melanog
12	557.5	6.9	20448	23 ABL19989	Drosophila melanog
13	537	6.7	24971	23 ABL19988	Drosophila melanog
14	530.5	6.6	24789	23 ABL28640	Drosophila melanog
15	492	6.1	7568	23 AAS80823	DNA encoding novel
16	492	6.1	9220	23 ABV24414	Human prostate exp
17	492	6.1	9220	23 ABV25265	Human prostate exp
18	492	6.1	9416	24 ABK83800	Human cDNA differe
19	490.5	6.1	10242	23 ABL20787	Drosophila melanog
20	462	5.7	5361	18 AAT78868	P. falciparum live
21	462	5.7	5529	24 ABK50883	DNA encoding plasm
22	462	5.7	6152	18 AAT78867	P. falciparum live
23	454	5.6	6919	23 ABL05593	Drosophila melanog
24	453.5	5.6	4407	23 ABL12349	Drosophila melanog
25	451	5.6	6561	23 ABL12348	Drosophila melanog
26	450.5	5.6	20978	23 ABL20786	Drosophila melanog
27	448	5.5	9953	23 ABL21371	Drosophila melanog
28	440.5	5.5	5439	23 ABL02411	Drosophila melanog
29	432	5.4	9373	23 ABL07021	Drosophila melanog
30	431.5	5.3	7997	22 ABN47417	Human breast cell
31	431.5	5.3	7997	22 ABN47417	Human foetal liver
32	431.5	5.3	7997	22 ABN47417	Probe #10872 for g
33	431.5	5.3	7997	22 ABN47417	Human brain expres
34	431.5	5.3	7997	22 ABN47417	Human bone marrow
35	431.5	5.3	7997	22 ABN47417	Probe #10203 for g
36	431.5	5.3	7997	22 ABN47417	Probe #14158 used t
37	431.5	5.3	7997	22 ABN47417	Human genome-deriv
38	431.5	5.3	7997	22 ABN47417	Human ORF ORF2255
39	431.5	5.3	11680	21 AAC76700	Drosophila melanog
40	430	5.3	12505	23 ABL07020	Drosophila melanog
41	429	5.3	8922	23 ABL04053	Drosophila melanog
42	429	5.3	15913	23 ABL04052	Drosophila melanog
43	428.5	5.3	6744	13 AAQ29471	Extracellular fact
44	427	5.3	15231	24 ABK63598	Rat sequence diffe
45	422	5.2	10253	24 ABL60871	Human cytoskeleton

ALIGNMENTS

RESULT 1	
AAV02302	
ID	AAV02302 standard; cDNA; 5200 BP.
XX	
AC	AAV02302;
XX	
DT	20-JUL-1998 (first entry)
XX	
DE	Rat tumour suppressor gene SSeCKS.
XX	
KW	SSeCKS; tumour suppressor gene; rat; protein kinase C; mitosis;
KW	cancer; malignancy; cell proliferation; Alzheimer's disease;
KW	therapy; ss.
XX	
OS	Rattus sp.
XX	
FT	Key Location/Qualifiers
	CDS 1..4791

```

FT      /+cag= a
XX      MO9740059-A1.
PN      30-OCT-1997.
PD      18-APR-1997; 97WO-US06830.
XX      18-JUN-1996; 96US-0665401.
PR      19-APR-1996; 96US-0635121.
XX      (GELM/) GELMAN I.
PA      (JAKE/) JAKEN S.
PI      Gelman I, Jaken S;
XX      WPI: 1997-535770/49.
DR      P-PSDB; AAM31347.
XX      Tumour suppressor gene SSeCKs - used as a mitotic regulator, and
PT      inhibitor of malignant phenotype
XX      Claim 1; Fig 11A-L; 162pp; English.
XX      This rat CDNA sequence codes for the full-length form of the novel
CC      tumour suppressor protein SSeCKs (see AAM31347). The SSeCKs gene
CC      product is a substrate of protein kinase C and acts as a negative
CC      regulator of mitosis and as an inhibitor of the transformed
CC      phenotype. A partial SSeCKs clone (see AAV02301) was identified by
CC      searching for cDNAs whose abundance was low in NIH 3T3 cells and
CC      decreased following the expression of the activated oncogene v-src.
CC      A 5'RACE product was spliced to this truncated sequence to construct
CC      the full-length SSeCKs cDNA sequence. This full-length SSeCKs
CC      nucleic acid sequence, as well as homologous and hybridising nucleic
CC      acids are claimed, as are vectors comprising such nucleic acids,
CC      encoded proteins, host cells and methods of inhibiting the
CC      expression of a transformed phenotype in a host cell by introducing
CC      the nucleic acid. Introduction of a SSeCKs nucleic acid or gene
CC      product into a host cell inhibits mitosis of the host cell,
CC      allowing the treatment of diseases associated with disorders of
CC      proliferation and/or with the expression of a malignant phenotype.
CC      SSeCKs can also be used to treat or identify disorders of
CC      cytoskeletal structure and cellular architecture (such as
CC      Alzheimer's disease), and may be a marker for aberrancies in
CC      fertility and/or nervous system development.
XX      Sequence 5200 BP; 1612 A; 1201 C; 1563 G; 824 T; 0 other;
SO
Alignment Scores:
Pred. No.: 0 Length: 5200
Score: 8041.00 Matches: 1592
Percent Similarity: 99.75% Conservative: 0
Best Local Similarity: 99.75% Mismatches: 4
Query Match: 99.60% Indels: 0
Gaps: 0
US-09-902-432-4 (1-1596) x AAV02302 (1-5200)
QY      1 MetGlyAlaGlySerSerThrGluGlnArgSerProGluGlnProAlaGlySerAspThr 20
Db      1 ATGGCGCGAGGCGATTCCACCGAGCAGCGAGCCCGGACACCGCGGGGAGCGACAG 60
QY      21 ProSerGluLeuValLeuSerGlyHisGlyProAlaAlaGluAlaSerGlyAlaGly 40
Db      61 CCGAGCGAGCTGTGTCTCAGTGGCCATGGGCCCGCAGCTGAAAGCTCGGAGCAGCTGGA 120
QY      41 AspProAlaAspAlaAspProAlaThrIlySLeuProGlnIlyAsnGlyIlyLeuSerSer 60
Db      121 GACCCGCGCGAGCGAGGCCCGCCACCAAGCTCCCAAGAAATGGCCAGCTGTCTTCT 180
QY      61 ValAsnGlyValAlaGluGlnGlyAspValHisValGlnGluGluAsnGlnGlnGlyGln 80
Db      181 GTCAACGGCGTAGTGAACAAGAGATGTCTCAAGAGGAAAAACAGAGAGGGGCGAG 240

```

```

QY      81 GluGluGluValValAspGluAspValGlyGlnArgGluSerGluAspValArgGlyIlys 100
Db      241 GAGGAAGAAGTCGTTGATGAGAGATGTTGGACGAGAGTCAGAAAGATGTGAGAGAAAAA 300
QY      101 AspArgValGluGluGluMetAlaAlaAsnSerThrAlaValGluAspIleThrIlyAspGly 120
Db      301 GACCGAGTTGAAGAAATGGCCGCAACTCCACAGCTGTGAAAGATATCAGAAAGGATGGG 360
QY      121 GlnGluGlnThrSerGluIleIleGluGlnIleProAlaSerGluAsnAsnValGluGlu 140
Db      361 CAGAGGAGACATCAGAAATTAATTGAACAGATCCCTGCTTCAGAAACAAATGTGGAAACA 420
QY      141 MetValGlnProAlaGluSerGlnAlaAsnAspValGlyPheIlySlysValPheIlySph 160
Db      421 ATGGTACAGCTGTGAGTCCAGGCTAATGATGTGGCTTCAAGAAAGTATTTAAATTT 480
QY      161 ValGlyPheIlySphThrValIlySlysAspIlySasnGlyIlySerAspThrValGluLeu 180
Db      481 GTTGGTTTATTTTCACGTTGAAGAGATATAAATGAAAAAGTCAGATATCTCCACTA 540
QY      181 LeuThrValIlySlysAspGluGlyGluGluValAlaGluAlaSerValGlyAlaGlyAspHis 200
Db      541 CTCACTGTCAAGAGAGATGAAGCGCAAGGCGCAGAGCTCTGTCTCGAGCTGGAGACAC 600
QY      201 GlnGluProSerValGluThrAlaValGlyGluSerAlaSerIlySlysGluSerGluLeuIlys 220
Db      601 CAGGAGCCCATGTGGAGACTCCGTCGAGAGTCAAGTCAAAAGAAATGAGCTGAAG 660
QY      221 GlnSerThrGlyIlySlysGlnGluGlyThrLeuIlySlysGlnGlnGlnSerSerThrGluIlePro 240
Db      661 CAATCCAGAGAAAGCAAGAAAGGACCTTGAAGCAAGAACAGAGACAGCAAGAAATCTCC 720
QY      241 LeuGlnAlaGluSerAspGlnAlaAlaGluGluGluAlaIlySAspGluGlyGluIlyIlys 260
Db      721 CTTCAAGCCGAATCTGATCAAGCCGCTAGAGAAAGCCAAAGATGAAGCAAGAAAAAA 780
QY      261 GlnGlyIlySlysGluProThrIlySerProGluSerProSerSerProValAsnSerGluThr 280
Db      781 CAAGAGAAAGAGCCCAACCAAGTCCCAAGATCCCGAGACACCCAGTCAACAGTAGACA 840
QY      281 ThrSerSerPheIlySlysPhePheThrHisGlyTTPAlaGlyTTPArgIlySlysThrSer 300
Db      841 ACATTTCTTCCAAATAATTTCTTCACTACGTTGGCGGCTGGCGAAGAGCCAGC 900
QY      301 PheIlySlysSerIlySlysAspAspLeuGlnThrAlaGluIlySArgIlyGlnGlnIlyAla 320
Db      901 TTCAAGAAATCAAAAGAGATGATCTGAAACTGCCGAAAGAGAAAGACCAAGAGCGCA 960
QY      321 GlnIlySValAspGluGluGluIlySlysGlyIlyThrGluProAlaSerGluGlnGlnPro 340
Db      961 GAAAAAGTAGACGAGGAAGAAAGAAAGACAGAGCCAGCTCGAGAGAGACAGAGCCG 1020
QY      341 AlaGluAspThrAspGlnAlaArgLeuSerAlaAspIlyIlyIlySValGluLeuProIleu 360
Db      1021 GCAGAAAGACACAGACCAAGCCAGCTGTTCACAGACTACAGAAAGTGAAGCTGCTTTG 1080
QY      361 GlnAspGlnValGlyAspLeuGlnAlaSerSerGluGluIlySlysCysAlaProLeuAlaThr 380
Db      1081 GAGACCAAGTTGTGATCTGAGGCAATCTCAGAGGAAAGATGTCTCTTTGGCAACG 1140
QY      381 GlnValPheAspGluIlyMetGlnAlaHisGlnGlnValValAlaGluValHisValSer 400
Db      1141 GAAGTGTTTGATGAAGAATGGAAGCCCAACAAAGAAATGTTGGAGAGTCCACCTGAGC 1200
QY      401 ThrValGluIlySThrGluGluGluGlnGlnIlyGlyGlyGlnAlaGluGlyValVal 420
Db      1201 ACCGTGAGAAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
QY      421 ValGluGlyThrGlyGluSerLeuProGluIlySLeuAlaGluProGlnGlnIlyValPro 440
Db      1261 GTAGAAAGAACAGAGAAATCTTGGCCCTTGAGAAACTGGCTGAGCCCAAGAGAGTCCCC 1320

```

QY 441 GlnGluAlaGluProAlaGluGluLeuMetLysSerArgGluMetCysValSerGlyGly 460
DB 1321 CAGGAAGCTGAGCTGCTGAGGAGCTGATGAAGACGACAGAGATGTGTCTCTGGAGGA 1380
QY 461 AspHisThrGlnLeuThrAspLeuSerProGluGluLysThrLeuProLysHisProGlu 480
DB 1381 GACCACACTCACTGACAGACCTAAGTCTCTGAAGAGACGCTGCCCAACACCCAGAA 1440
QY 481 GlyIleValSerGluValGluMetLeuSerSerGlnGluArgIleLysValGlnGlySer 500
DB 1441 GGCATTGTCTAGTGGTGGAGATGCTCTCTCTCAGGAAAGAAATCAAGGTACAGGGAAGT 1500
QY 501 ProLeuLysLysLeuPheSerSerSerGlyLeuLysLysLeuSerGlyLysLysGlnLys 520
DB 1501 CCCTTGAAGAAATCTTTCAGTAGCTCAGGCTTAAAGAAAGCTGTCTGGAAAGACGAGAAG 1560
QY 521 GlyLysArgGlyGlyGlyAspGluGluProGlyGluPyrGlnHisIleHisThrGlu 540
DB 1561 GGGAAACGAGGAGGTGGGGAGACGAAGAGCTCTGGAGATACCACACATTCACACCGAA 1620
QY 541 SerProGluSerAlaAspGluGlnLysGlyLysSerSerAlaSerSerProGluGluPro 560
DB 1621 TCCCCAGAGAGTGTGATGACGACGAAGGAGAGAGCTCTCGTCTGCCCGAGGAGCCT 1680
QY 561 GluGluThrThrCysLeuGluLysGlyProLeuGluAlaProGlnAspGlyGluAlaGlu 580
DB 1681 GAGGAGACCACTGTCTGGAGAAAGGCCCTGGAAAGCACCCAGGATGGGGAAGCTGAG 1740
QY 581 GluGlyThrThrSerAspGlyGluLysLysArgGluGlyIleThrProTppAlaSerPhe 600
DB 1741 GAAGGAACACTTCCGATGGAGAGAAAGAGAGAGAGGATCACTCCCTGGGCATCCTTC 1800
QY 601 LysLysMetValThrProLysLysArgValArgProSerGluSerAspLysGluGlu 620
DB 1801 AAAAGATGTGTACACCCAAAGAAACGGTCCGAAGACTTCTCGAGAGTGCACAGGAGGAA 1860
QY 621 GluLeuGluLysValLysSerAlaThrLeuSerSerThrAspSerThrValSerGluMet 640
DB 1861 GAGCTGGAGAAAGTCAAGACGCCACCTTGTCTCTCACTATAGCACAGTGTCAAGAAATG 1920
QY 641 GlnAspGluValLysThrValGlyGluGluGlnLysProGluGluProLysArgArgVal 660
DB 1921 CAAGATGAAGTCAAACTGTGTGAGGAACAAAGCCAGAGGAAACCAAGCGTAGGGTG 1980
QY 661 AspThrSerValSerTppGluAlaLeuIleCysValGlySerSerLysLysArgAlaArg 680
DB 1981 GATACTTCAGTGTCTTGGGAAGCACTGATTGTGTCTCGATCATCCAAAGAGAGAGCAAGG 2040
QY 681 LysAlaSerSerSerAspAspGluGlyGlyProArgThrLeuGlyGlyAspSerHisArg 700
DB 2041 AAGGCATCTCTTTCAGATGATGAAGGAGGGCCAAAGGACACTGGGAGGGGACAGTCAAGA 2100
QY 701 AlaGluGluAlaSerLysAspLysGluAlaGlyThrAspAlaValProAlaSerThrGln 720
DB 2101 GCAGAGAGGCCAGCAAGAACAAGAAAGCCGGAACAGACGCTGTCTCTGCGACCCACAG 2160
QY 721 GluGlnAspGlnAlaGlnGlySerSerProGluProAlaGlySerProSerGluGly 740
DB 2161 GAGCAGGACCAAGCGCAAGGAAGTCTCTACCCGAGCCAGCGGGAAGCCCTTCCGAAGGG 2220
QY 741 GluGlyValSerThrTppGluSerPheLysArgLeuValThrProArgLysLysSerLys 760
DB 2221 GAAGGTCTCTCCACTTGGGAGTCAATTTAAAGATTAGTCACTCCACAGAAAAAATCCAAAG 2280
QY 761 SerLysLeuGluGluLysAlaGluAspSerSerValGluGlnLeuSerThrGluIleGlu 780
DB 2281 TCAAACTGGAAGAGAAAGCCGAAGACTCTAGTGTAGAGCAGTTGTCCACTGAGATCGAA 2340
QY 781 ProSerArgGluGluSerTppValSerIleLysLysPheIleProGlyArgArgLysLys 800
DB 2341 CCGAGTAGAGAGAAATCTTGGGTCTTCCATTAAAGAAATTCATCCCGGACGGGGAAGAAA 2400
QY 801 ArgAlaAspGlyLysGlnGluGlnAlaThrValGluAspSerGlyProValGluIleAsn 820

DB 2401 AGGCGACAGCGGGAAGCAAGAACCAAGCCACTGTGGAAGACTCAGGGCCAGTGGAGATAAAT 2460
QY 821 GluAspAspProAsnValProAlaValValProLeuSerGluTyrAsnAlaValGluArg 840
DB 2461 GAGGACGACCCCTAATGTCTCCAGCCGTCGTCTCTGAGTATAATGTCAGTGGAGAGG 2520
QY 841 GluLysMetGluAlaGlnGlyAsnThrGluLeuProGlnLeuLeuGlyAlaValTyrVal 860
DB 2521 GAGAAGATGGAAGCCAGGGGAATACGGAGCTGCCCCAGCTGCTGGGGCTGTGTACGTG 2580
QY 861 SerGluGluLeuSerLysThrLeuValHisThrValSerValAlaValIleAspGlyThr 880
DB 2581 TCCGAGGAGCTCAGTAAGACTCTGGTCCACACTGTGAGTGTCCAGTCAATGATGGAGCC 2640
QY 881 ArgAlaValThrSerValGluGluArgSerProSerTrpIleSerAlaSerValThrGlu 900
DB 2641 AGGCGAGTCACCACTGCTGAAGAGCGGTCTCTCTCTGGATATCCGCTCCGTAACAGAA 2700
QY 901 ProLeuGluHisThrAlaGlyGluAlaMetProProValGluGluValThrGluLysAsp 920
DB 2701 CCTCTTGAACACACACAGCGGGAAGCCATGCCACCTGTTTGAAGAGGTCACTGAAAAAGAC 2760
QY 921 IleIleAlaGluGluThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHis 940
DB 2761 ATCATTTGCAGAAAGAACTCTCTGTCTCACCCAGACGTTACCAGAGGTAAAGATGCCCAT 2820
QY 941 AspAspMetValThrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThr 960
DB 2821 GACGACATGGTCACCACTGAAGTGGATTTACCTCAGAGCTGTGACAGCCACAGAGACC 2880
QY 961 SerGluAlaLeuArgThrGluGluValThrGluAlaSerGlyAlaGluGluThrThrAsp 980
DB 2881 TCAGAGGCTCTCCGTACTCGAAGAGTTTACCGAAGCATCGGGGGCCGAAGAGACACACAGAC 2940
QY 981 MetValSerAlaValSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrPro 1000
DB 2941 ATGGTGTCCGACGTTTCCACAGCTGACTGCTCCACAGACACACAGAGGAAGGCCACCCCA 3000
QY 1001 ValGlnGluValGluSerGlyValLeuAspThrGluGluGluArgGlnThrGlnAla 1020
DB 3001 GTTCAGAGGTAGAGAGTGGTGTCTAGATACAGAGAGAGAGAGCGCCACAGCGAGGCC 3060
QY 1021 IleLeuGlnAlaValAlaAspLysValLysGluLysSerGlnValProAlaThrGlnThr 1040
DB 3061 ATCTCTCAAGCCGTTCGAGACAAGGTGAAGAGAGTCCAGGTGCTGCAACCCAGACT 3120
QY 1041 ValGlnArgThrGlySerLysAlaLeuGluLysValGluGluValGluGluAspSerGlu 1060
DB 3121 GTGCAGAGAACCGGGTCAAAGACACTGGAGAGGTTGAGAGAGGTAGAGAGGACTCCGAA 3180
QY 1061 ValLeuAlaSerGluLysGluLysAspValMetProLysGlyProValGlnGluAlaGly 1080
DB 3181 GTCTGTGCTTCGAGAGAAAGAGAGGACCTTATGCCGGAAGACCCGTGCGAGGAAGCTGGA 3240
QY 1081 AlaGluHisLeuAlaGlnGlySerGluThrGlyGlnAlaThrProGluSerLeuGluVal 1100
DB 3241 GCTGAGCATCTTCACAGGGCTCTGAGACTGGACAGGCTACTCCAGAGAGCCCTTGAAGTT 3300
QY 1101 ProGluValThrAlaAspValAspHisValAlaThrCysGlnValIleLysLeuGlnGln 1120
DB 3301 CCTGAAGTCAAGCCAGATGTAGACCATGTGCCACGCTGCCAGGTTTCAAGCTCCAGCAG 3360
QY 1121 LeuMetGluGlnAlaValAlaProGluSerSerGluThrLeuThrAspSerGluThrAsn 1140
DB 3361 CTGATGGAAACAGGCCCTGGGCCCTGAGTCACTCCGAAACCTTGACAGACAGTGGAGCAAT 3420
QY 1141 GlySerThrProLeuAlaAspSerAspThrAlaAspGlyThrGlnGlnAspGluThrIle 1160
DB 3421 GGAAGCACTCCCTTAGCAGATTACAGACTGCAGACTGGGACACAGCAAGATGAATCAAT 3480
QY 1161 AspSerGlnAspSerLysAlaThrAlaValArgGlnSerGlnValThrGluGluGlu 1180

D	b	3481	GAGCGCCAGACAGTAAAGCCACTGCAGCTGTCAGGCACTACAGTCAACAGAAAGAG	3540
Q	y	1181	AlaAlaThrAlaGlnIlysgIugIupProSerThrLeuProAsnValProAlGlnIu	1200
D	b	3541	GGCGGTACTGTCGAAAGAGGAGGACCTTGACACTCACTAATATATGTTCCAGCCAGGAA	3600
Q	y	1201	GIuHIsGlyIugIuPProGlyIaIgaAspValLeuGluPProThGInGInGluLeuThrAla	1220
D	b	3601	GAACATGGGGAGAACACAGAGAGATGTTCTTGAACCTTACACACCAAGACTTACTGCT	3660
Q	y	1221	AlaAlaValProValLeuValAlaYsThrGluValAlGlyGluGlyGlyValAspTyrLeu	1240
D	b	3661	GCAGCCGTGCCCCCTTCTTGGCCAAAGACTGAGGTGGCTCAAGAGGGTGAAGCTTGCTG	3720
Q	y	1241	AspGlyIugIyValIySgIugIugInGluValPheValHisSerGlyProAsnSerGln	1260
D	b	3721	GATGGAGAAAAGTCAAAAGAAACAGAGAGTGTGTTGTACACTTGGACCCAAAGTCA	3780
Q	y	1261	LysAlaAlaAspValThrTyrAspSerGluValIleMetClyValAlGlyCySGInGluIyS	1280
D	b	3781	AAGCGTGTGATGTGACATATGACAGTGAAGTATGGAGTGGCGGGGTGTACAGAAAAG	3840
Q	y	1281	GIuSerThrGluValAlGInSerLeuSerLeuGlyGlyGlyIleMetGlyIuThrAspValGlu	1300
D	b	3841	GAGACTACTGAAGTGCAGAGACTTTAGCTCTGGAGAGGAGAGATGAAACTGACCTTGA	3900
Q	y	1301	LysGlyIySArGgIuThrIySProGluGlnValSerGlyGluGlyGlyGlnGluThrAla	1320
D	b	3901	AAGGAGAAAAGGAGACAAAGCCAGACAGACAGTGAAGAAAGTGAAGAGAAACAGCC	3960
Q	y	1321	AlaProGluHIsGlyIyGlyThrTyrGlyIySProValLeuThrLeuAspMetProSerSer	1340
D	b	3961	GCTCTGTGACATGAAGAACCTTCAGGGAAACCAGTCTTCACTGACATGCCAGCTCA	4020
Q	y	1341	GIuArGgIyLysValAlaGlnSerLeuGlyGlyIleSerProSerLeuProAspGluAspIyS	1360
D	b	4021	GAGAGGGGGAAGGACATGGGAGGACTTGGAGAGGCCCTTCTCTCCAGACCAAGACAA	4080
Q	y	1361	AlaGlyCySIIeGlyValGlnValGlnSerLeuAspThrThrValThrGlnThrAlaGlu	1380
D	b	4081	GCAGTTGCCATAGAGGTTTCAAGTTCAAAAGCTGTGACACAAACAGTCACTCAAAACAGCAA	4140
Q	y	1381	AlaValGlyIySValIIeGlyuThrValValIIeSerGlyuThrGlyIyGluSerProGluCyS	1400
D	b	4141	GCTGTGAAAAGGTGCATAGAAACCGTTGTATTTCAAGAGCAGGTGAAGGTCCAGACGT	4200
Q	y	1401	ValGlyAlaHIsleuLeuProAlaGlnIySLeuSerAlaThrGlyGlyHIsTyrThrLeu	1420
D	b	4201	GTAGGTGACCACTTATTCACGCTGAGAAATCTCTTGCACCGGTGGCCACTGCACTTT	4260
Q	y	1421	GlnHisAlaGluAspThrValProLeuGlyProGluSerGlnAlaGluSerIIeProIIe	1440
D	b	4261	CAGCATGCAGAGGACAGGTATCCCTGGGGCCGTGAAGTCTCAGGCAAGAAATCCATCCAAATC	4320
Q	y	1441	IleValIThrProAlaProGluSerThrLeuHisProAspLeuGlnGlyuIIeSerAla	1460
D	b	4321	ATAGTAACTCTCTGCTCTGAAAGCACCTTCATCTCGAACCCTCAAGAGAAATAAGCCCA	4380
Q	y	1461	SerGlnaGgIyuaGysSerGluGlnGluAspIySProAspAlGlyProAspAlaAspGly	1480
D	b	4381	TCCCAAGAAAGGACATCAAGAGAAAGAGACACAGCCAGATCTGGTCTGATCTCAAGCC	4440
Q	y	1481	LysGluSerThrAlaIIeGlyIySValLeuIyAlaGluProGluuIIeGluIuGlu	1500
D	b	4441	AAAGGAGTTACAGCATTCGAAAAGTCTCAAGGCTGAAGCTGAGATCCTGGAATCTTGAG	4500
Q	y	1501	SerIySArAsnIySIIeValIleuAsnValIIeGlnThrAlaValaAspGlnPheAlaArg	1520
D	b	4501	AGTAAAGACCAACAAGATGTGCTGAACGTATTCACAGACAGCGTGAACAGTTCCAGACT	4560
Q	y	1521	ThrGluThrAlaProGluThrHisAlaIlyAspSerGlnThrGlnValProAlaGlyArg	1540
D	b	4561	ACAGAAACAGCCCCGAAACTCAAGCTTAATGATTCACAGCCCAAGGATTCCTGACGACGG	4620

Accession	Gene	Protein	Location/Qualifiers
Oy	1541	LeuaspserargGUP-roAaenargCysTPThrIysMetLysAspAlaLysMetLysHis	1560
Db	4621	CTTGACGACGGGGAGCCCAACAGATCTGTGACAAAAATGAAAGAGCCAGATGAACAC	4680
Oy	1561	ProValProGlnProArgGlnuAspLeuGlnValLeuThrValLeuGluAlaTPAlaGln	1580
Db	4661	CCATGCGCGACGCCACAGAGAGACCTTGCAAGTCTGACCGCTTCTGGAGGATGGGCTCAG	4740
Oy	1581	ProArgLysCysLeuProArgLeuGlnLeuLysAlaProValSerLys	1596
Db	4741	CCTGGGAAATGCTTGCCGCGCTTGACAGTTGAAAGGCCCGGTGTCAAG	4788
RESULT 2			
Id	AAV02301	standard; cDNA, 5074 BP.	
XX	AAV02301;		
AC	AAV02301;		
XX			
DT	20-JUL-1998	(first entry)	
XX			
DE	Rat tumour suppressor gene SSeCKs (truncated form).		
XX			
KW	SSeCKs; tumour suppressor gene; rat; protein kinase C; mitosis;		
KW	cancer; malignancy; cell proliferation; Alzheimer's disease;		
KW	therapy; ss.		
OS	Rattus sp.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	176..4216	
FT		/*rag= a	
FT		/transl_except= (pos:189..391, aa:Glu)	
FT		/transl_except= (pos:1661..1663, aa:Lys)	
FT		/transl_except= (pos:1668..1670, aa:Lys)	
FT		/transl_except= (pos:2444..2446, aa:Arg)	
FT		/transl_except= (pos:2528..2530, aa:Glu)	
FT	polyA_signal	2528..2530	
FT		/*rag= b	
FT	polyA_signal	5056..5061	
FT		/*rag= c	
XX			
PN	WO9740059-A1.		
PD	30-OCT-1997.		
XX			
PF	18-APR-1997;	97MO-US06830.	
XX			
PR	18-JUN-1996;	96US-0665401.	
PR	19-APR-1996;	96US-0635121.	
XX			
PA	(GELM/) GELMAN I.		
XX	(JAKE/) JAKEN S.		
PI	Gelman I, Jaken S;		
XX			
DR	WPI, 1997-535770/49.		
XX	P-PEDB; AAM31346.		
PT	Tumour suppressor gene SSeCKs - used as a mitotic regulator, and		
XX	inhibitor of malignant phenotype		
XX			
XX	Example 6; Fig 3A-H, 162p; English.		
CC	This rat cDNA sequence codes for an active truncated form of the novel tumour suppressor protein SSeCKs (see AAM31346). The SSeCKs gene product is a substrate of protein kinase C and acts as a negative regulator of mitosis and as an inhibitor of the transformed phenotype. The SSeCKs clone was identified by searching for cDNAs whose abundance was low in NIH 3T3 cells and decreased following the expression of the activated oncogene v-src. A 5'RACE product was spliced to the truncated sequence to construct a full-length SSeCKs cDNA (see AAV02302). This full-length SSeCKs nucleic acid		

PS Example 6; Fig 3A-H; 162pp; English.

CC sequence, as well as homologous and hybridising nucleic acids, are
 CC claimed, as are vectors comprising such nucleic acids, host cells
 CC and methods of inhibiting the expression of a transformed phenotype
 CC in a host cell by introducing such nucleic acids. Introduction of
 CC a SSeCKS nucleic acid or gene product into a host cell inhibits
 CC mitosis of the host cell, allowing the treatment of diseases
 CC associated with disorders of proliferation and/or with the
 CC expression of a malignant phenotype. SSeCKS can also be used to
 CC treat or identify disorders of cytoskeletal structure and cellular
 CC architecture (such as Alzheimer's disease), and may be a marker
 CC for aberrancies in fertility and/or nervous system development.
 XX
 SQ Sequence 5074 BP; 1501 A; 1151 C; 1441 G; 981 T; 0 other;

Alignment Scores:

Pred. No.: 3,15e-300 Length: 5074
 Score: 6001.00 Matches: 1231
 Percent Similarity: 97.33% Conservatives: 8
 Best Local Similarity: 96.70% Mismatches: 27
 Query Match: 74.33% Indels: 14
 DB: 18 Gaps: 4

US-09-902-432-4 (1-1596) x AAV02301 (1-5074)

QY 329 GluLysThrGluProAlaSerGluGluGlnGluProAlaGluAspThrAspGlnAlaArg 348
 Db 2 GAAAGACAGAGCCAGCTCGGAGGAGCAGAGCCGCGAAGACACAGACCAGGCCAGG 61
 QY 349 LeuSerAlaAspTyrGluLysValGluLeuProLeuGluAspGlnValGlyAspLeuGlu 368
 Db 62 TTGTCACGACACTACGAGAGAGGTGGAGCTGCCTTTGGAACCCAGGTTGGTGACCTGGAG 121
 QY 369 AlaSerSerGluGluLysCysAlaProLeuAlaThrGluValPheAspGluLysMetGlu 388
 Db 122 GCATCGTCAGAGGAGAAGTGTGCTCTTGGCAACGAGAGTGTGATGAGAGATGGAA 181
 QY 389 AlaHisGlnGluValAlaGluValHisValSerThrValGluLysThrGluGluGlu 408
 Db 182 GCCCACCAAGAGTGTGTGAGAGGTCACGTCGAGCACCGTGGAGAACAGACAGAGGAG 241
 QY 409 GlnGlyGlyGlyGluAlaGluGlyValValGluValGluGlyThrGlyGluSerLeu 428
 Db 242 CAGGAGGAGGAGAGGCTGAAGGGGGGCTGGTGTGTAGAGGAACAGAGAGATCCTTG 301
 QY 429 ProProGluLysLeuAlaGluProGlnGluValProGlnGluAlaGluProAlaGluGlu 448
 Db 302 CCCCTCAGAGAACTGGCTGAGCCCGCAGAGGTCCCCCAGGAAGCTGAGCCTGCTGAGGAG 361
 QY 449 LeuMetLysSerArgGluMetCysValSerGlyGlyAspHisThrGlnLeuThrAspLeu 468
 Db 362 CTGATGAAGAGCAGAGAGATGTGTGCTCTGGAGGAGACACACTCAACTGACAGACCTA 421
 QY 469 SerProGluGluLysThrLeuProLysHisProGluGlyLeuValSerGluValGluMet 488
 Db 422 AGTCCTGAAGAGAGACGCTGCCCAACACCCAGAGGCAATTGTGAGTGAAGTGGAGATG 481
 QY 489 LeuSerSerGlnGluArgIleLysValGlnGlySerProLeuLysLysLeuPheSerSer 508
 Db 482 CTGTCCTCTCAGGAAGAAATCAAGGTACAGGGAAGTCCCTTTGAAGAACTCTTCAGTAGC 541
 QY 509 SerGlyLeuLysLysLeuSerGlyLysLysGlnLysGlyLysArgGlyGlyGlyAsp 528
 Db 542 TCAGGCTTAAAGAGCTGTCTGGAGAGAGCAGAGGGAACGAGAGGCTGGGGGAGAC 601
 QY 529 GluGluProGlyGluTyrGlnHisIleHisThrGluSerProGluSerAlaAspGluGln 548
 Db 602 CAAGAGCCTCGAGAAATACCAACACATTCAACCCGAATCCCCAGAGAGTGTGTATGAGCAG 661
 QY 549 LysGlyGluSerSerAlaSerSerProGluGluProGluGluThrThrCysLeuGluLys 568
 Db 662 AAGGGAGAGAGCTCTGGTCTGCCCGAGAGGCTGAGAGAGACCAGCTGTCTGGAGAA 721
 QY 569 GlyProLeuGluAlaProGlnAspGlyGluAlaGluGlyThrThrSerAspGlyGlu 588

Db 722 GGCCCGCTGAAGACACCCAG-GATGGGAGAGCTGAGAAAGAACTACTTCGT--GGAGAG 778
 QY 589 LysLysArgGluGlyIleThrProTyrAlaSerPheLysLysMetValThrProLysLys 608
 Db 779 AAGAAGAGGAAGG--ATCAGCTCCTGGGCATCCTTCAAAAAAGATGGTGACACCCNAGAAA 836
 QY 609 ArgValAlaArgProSerGluSerAspLysGluGluGluLeuGluLysValLysSerAla 628
 Db 837 CGGTC-CGAAGACCTTCTGAGAGTCAAGAGGAGAAAGCTGGAGAGGCTCAAGAGCGCC 895
 QY 629 ThrLeuSerSerThrAspSerThrValSerGluMetGlnAspGluValLysThrValGly 648
 Db 896 ACCTTGTCTCCTCACTGATAGCAGGTGCAGAAATCAAGATGAAGTCAAACTGTGTGT 955
 QY 649 GluGluGlnLysProGluGluProLysArgValAspThrSerValSerTyrGluAla 668
 Db 956 GAGGAACAAAAAGCCAGAGGAACCAAGAGCGTAGGCTGGATACTTCAGTGTCTTGGGAAGCA 1015
 QY 669 LeuIleCysValGlySerSerLysLysArgAlaArgLysAlaSerSerSerAspAspGlu 688
 Db 1016 CTGATTTGTGTCGGATCATCCAAAGAGAGAGCAAGGAAGGCATCTCTTCAGATATAAGA 1075
 QY 689 GlyGlyProArgThrLeuGly-GlyAspSerHisArgAlaGluGluAlaSerLysAspLys 708
 Db 1076 ---GGGCCAAGGACACTGGGAGGGGACAGTCAACAGCAGAGAGGAGGCCAGCAAGACAA 1132
 QY 708 sGluAlaGlyThrAspAlaValProAlaSerThrGlnGluGlnAspGlnAlaGlnGlySe 728
 Db 1133 AGAAGCGGA-ACAGAGCGCTGTTCTGCCAGCACCCAGGAGCAGACCAAGCGCAGAGAG 1191
 QY 728 rSerSerProGluProAlaGlySerProSerGluGlyGluGlyValSerThrTrpGluSe 748
 Db 1192 TTCCTCACCGAGCCAGCGGGAAGCCCTCCGAAGGGGAAGGTGTCTCCACTTTGGGAGTC 1251
 QY 748 rPheLysArgLeuValThrProArgLysLysSerLysSerLysLeuGluLysAla-- 767
 Db 1252 ATTTAAAGATTAGTCACCTCAAGAAAAAATCCAAAGTCAAACTGGAAGAGAAAGA-AG 1310
 QY 768 ---GluAspSerSerVal--GluGlnLeuSerThrGluIleGluProSerArgGluGl 785
 Db 1311 CCGGAGGAGACTCTAGTGTGTAGGAGCAGGTGTCCACTGAGATCGAACCCGTGTAGAGAGA 1370
 QY 785 uSerTrpValSerIleLysLysPheIleProGlyArgArgLysLysArgAlaAspGlyLys 805
 Db 1371 ATCTTGGGTTTCCATTAAAGAAATTCATCCCGAGCGCGGAAGAAAGGSCACATGSGAA 1430
 QY 805 s-GlnGlnGlnAlaThrValGluAspSerGlyProValGluIleAsnGluAspAspPro 825
 Db 1431 GGCAAGAACCAAGCCACTGTGGAAGACTCAGGGCCAGTGGAGATAAATGAGGACGAGCCTG 1490
 QY 825 snValProAlaValValProLeuSerGluTyrAsnAlaValGluArgGluLysMetGlu 845
 Db 1491 ATGTCACAGAGCTGTGTCCTCTGTGATGATGATGATGATGATGATGATGATGATGATG 1550
 QY 845 laGlnGlyAsnThrGluLeuProGlnLeuGlyAlaValTyrValSerGluGluLeuS 865
 Db 1551 CCCAGGGGAATGCGGAGCTGCCAGCTGTGGGGCTGTGTA---GTGTCGAGAGAGCTCA 1607
 QY 865 erLysThrLeuValHisThrValSerValAlaValIleAspGlyThrArgAlaValThrS 885
 Db 1608 GTAAGACTCTGGTCCACACTGTGAGTGTGCGAGTCAATTGATGGACAGGAGGAGTCA 1667
 QY 885 erValGluGluArgSerProSerTrpIleSerAlaSerValThrGluProLeuGluHist 905
 Db 1668 GTCTCAAGAGGCGGTCTCTCTGTGATATCCGCTTCCGTTAAACAGAACCTCTTTGAACACA 1727
 QY 905 hrAlaGlyGluAlaMetProProValGluGluValThrGluLysAspIleIleAlaGluG 925
 Db 1728 CAGCGGGAGAGCCATGCCACCTGTTGAAGAGGTCACTGAAAAAGACATCATTTGCGAAG 1787
 QY 925 luThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHisAspAspMetValT 945

Db 1788 AAACCTCTGTGTCACCCAGACGGTTACCGAGGTCGAAGATGCCCATGACGATGCTCA 1847
 QY hrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeuA 965
 Db 1848 CCAGTGAAGTGGATTTCACCTCAGAAAGCTGTGACAGCCACAGAGACCTCAGAGGCTCTCC 1907
 QY rgThrGluGluValThrGluAlaSerGluValaGluGluThrThrAspMetValSerAlaV 985
 Db 1908 GTACTGAAGAATTACCGAAGCATCGGGGGCCGAAGAAGCCACAGACATGATGTCGGCAG 1967
 QY aLserGluLeuThrAspSerProAspThrThrGluGluValaThrProValaGluGluValG 1005
 Db 1968 TTTCCACGCTGACTGACTCCAGACACACAGAGSAAACCCAGTTCCAGAGGTG 2027
 QY luserGlyValLeuAspThrGluGluGluValaGluValaGluThrGluAlaLeuGluAlaV 1025
 Db 2028 AGGGTGTGTCTAGATACAGAAAGAGAGAGCCGACAGCAGCGACCTCTCAAGCCG 2087
 QY aAlaAspValValaGluGluGluSerGluValaProAlaThrGluThrValaGluThrG 1045
 Db 2088 TTGCAGACAAAGGTGAAGAAGAGAGTCCAGGTGCTGCAACCCAGACTGTGACAGAAACGG 2147
 QY lYserTyrAlaLeuGluValaGluValaGluValaGluValaGluValaGluValaGluVala 1065
 Db 2148 GGTCAAAAGCACTGAGAGAAAGTTGAGAGGTGAGAGAGGACTCCGAAAGTGTGCTTCGG 2207
 QY lulysGluValAspValaMetProLysGlyProValaGluGluValaGluValaGluVala 1085
 Db 2208 AGAAAGAGAGAGAGCTTATGCGAAAGAGACCCGTGACAGAACTGGAGCTGAGCATCTTG 2267
 QY laGluGluSerGluThrGluGluValaThrProGluSerLeuGluValaProGluValaThrA 1105
 Db 2268 CACAGAGGCTCTAGAGCTGAGCAGGCTACTCCAGAGAGCCCTTGAAGTCTCTGAAGTACAG 2327
 QY laAspValaAspHisValaAlaThrCysGluValaLeuLysLeuGluGluGluGluGluGlu 1125
 Db 2328 CAGATGTAGACCAATGTCCGACAGTGCAGGTTATCAAGCTCCAGACGCTGATGGAAACAG 2387
 QY lValAlaProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrProL 1145
 Db 2388 CCGTGGCCCTGAGTCATCCGAAACCTTGACAGACAGTGAACAATGGAGAACACTCCCT 2447
 QY eulAlaAspSerAspThrAlaAspGlyThrGluGluGluGluGluGluGluGluGluGluGlu 1165
 Db 2448 TAGCAGATTCAACAACCTGAGATGGGACACAGCAAGATGAAACCATGTGACCCAGAGACA 2507
 QY eLysAlaAlaThrAlaAlaValaArgInserGluValaThrGluGluGluGluGluGluGlu 1185
 Db 2508 GTAAAGCCACTGAGCTGTCAAGGCACTCAGGTCCAGAGTCCAGAAAGAGCGGCTACTGCTC 2567
 QY lulysGluGluProSerThrLeuProAsnAsnValaProAlaGluGluGluGluGluGluGlu 1205
 Db 2568 AGAAAGAGAGGCTTCGACACTAATTAATGTTCCAGCCCGAAGAACTAGGGGAAG 2627
 QY lProGluValaAspValaLeuGluProThrGluGluGluGluGluGluGluGluGluGluGlu 1225
 Db 2628 AACCGAAGAGATGTTCTTGAACCTCAACAGCAAGAGCTTGTCTGAGCCGTCGCG 2687
 QY aLeuAlaLysThrGluValaGluGluGluGluGluValaAspThrLeuAspGluGluLysV 1245
 Db 2688 TTGGCAAAAGACTGAGGTGGGTCAAGAGGCTGAGGTTACTGTGTTGATGAGAAAG 2747
 QY aLysGluGluGluGluGluValaPheValaHisSerGlyProAsnSerGluValaAlaAspV 1265
 Db 2748 TCAAGAAAGAACAGAGGAGGTTTGTACTCTGACCAACAGTCAAAAGGCTGCTGATG 2807
 QY aLThrTyrAspSerGluValaMetGlyValaAlaGlyCysGluGluGluGluGluGluGlu 1285
 Db 2808 TGACATATATGACGTGAAGATGAGGTGGCCGGGTGTAGAGAAAGAGAGACTGAAG 2867
 QY aLGlSerLeuSerLeuGluGluGluGluMetGluThrAspValaGluLysGluLysArg 1305
 Db 2868 TGCAAGTCTTAGCTCGAGAGGAGAGATGGAACCTGACGTTGAAGAGAGAAAGGG 2927

QY 1305 lUThrLysProGluGluValaSerGluGluGluGluGluGluGluGluGluGluGluGluGlu 1325
 Db 2928 AGACAAAGCCAGAGCAAGTGTAGTGAAGAAGTGAAGCAAGAAACAGCCCTCTGAGCATG 2987
 QY lUglYThrTyrGlyLysProValLeuThrLeuAspMetProSerSerGluArgGlyLysA 1345
 Db 2988 AAAGAACTACGGGAAGCCAGTCTCTGACACTTGACATGCCACAGCTCAGAGAGGGGAAG 3047
 QY lAlaGluSerLeuGlyLysProSerLeuProAspGluAspLysAlaGlyCysIleG 1365
 Db 3048 CACTGGAAAGCCTTGAGAGAGCCCTCTCTCCAGACCAAGACAAAGACAGTTGATG 3107
 QY lUValGluValaGluSerLeuAspThrThrValaThrGluThrAlaGluAlaValaGluLysV 1385
 Db 3108 AGGTTCAAGTTCAAAGCCTTGACACACAGCATCTCAAAAGCAGAAAGCTGTGAAAAG 3167
 QY aLlLeGluThrValValaLlSerGluThrGluGluGluSerProGluCysValaGluAlaHis 1405
 Db 3168 TCATGAAACGGTGTGTGATTTCAAGACAGGTGAAGTCCAGAGTGTAGTGCACACT 3227
 QY eulLeuProAlaGluLysSerSerAlaThrGlyHisTyrThrLeuGluHisAlaGluA 1425
 Db 3228 TATTACAGCTGAGAAAGTCTCTGCAACGGGTGGCCACTGACCTTTCAGCATGACAG 3287
 QY sPThrValProLeuGluProGluSerGluAlaGluSerIleProIleIleValaThrProA 1445
 Db 3288 ACACGGTACCCCTGGGGCTTCAAGTCTCAGGACAGATCCATCCCATCTAGTAACCTCTG 3347
 QY lAProGluSerThrLeuHisProAspLeuGluGluGluIleSerAlaSerGluArgGluA 1465
 Db 3348 CTCTGAAAGACCTTAATCTTAACCTCAAGAGAAATTAAGGCATCCAGAGAGAGC 3407
 QY rGSerGluGluGluValaAspLysProAspAlaGlyProAspAlaAspGlyLysGluSerThra 1485
 Db 3408 GATCAGAGAGAGAGCAAGCAAGTGTGCTGATCTGATGTCGAGCGCAAGAGAGATACAG 3467
 QY lAlleGluLysValLeuLysAlaGluProGluLlLeuGluGluGluGluGluGluGluGluGlu 1505
 Db 3468 CAATGCACAAAGTCTCTCAAGCTGAACCTGAGATCCTGGAACCTTGAGAGTAAAGCAACA 3527
 QY ySileValLeuAsnValaLlLeuGluThrAlaValaAspGluPheAlaArgThrGluThrAlaP 1525
 Db 3528 AGATTGTCTGACGTCATTCAAGACGCCGTTGACCAAGTTCGACATGACAGAAACAGCCC 3587
 QY roGluThrHisAlaTyrAspSerGluThrGluValaProAlaCysArgLeuAspSerArg 1545
 Db 3588 CCGAAACTCAATGCTTATGATTCACAGACCCAGGTTCTGCAATGGCTTGAGACAGAGG 3647
 QY lUProAsnArgCysTyrThrLysMetLysAspAlaLysMetLysHisProValProGluNP 1565
 Db 3648 AGCCCAACAGATGCTGGAACAAAGTGAAGTGGCCCAAGATGAACACCCAGTCCGAGC 3707
 QY roArgGluAspLeuGluValaLeuThrValaLeuGluAlaThrAlaGluProArgLysCysL 1585
 Db 3708 CCAGAGAGACTTGCAAGTCTGACCGTTCTGAGAGCAATGC--TCAGCTGGAAATGCT 3765
 QY l585 euProArgLeuGluLeuLysAlaProValaSerLys 1596
 Db 3766 TGCCGCGCTTGAGATTGAAGGCGCGGTGTCAAG 3800

RESULT 3

ABN59861 standard; cDNA; 6614 BP.

ABN59861;

28-JUN-2002 (first entry)

Novel human coding sequence SEQ ID NO: 272.

Human; anti-inflammatory; vulnerability; anti-inflammation; immunomodulator; anti-inflammation; cerebroprotective; cytoskeletal; rheumatic; gene therapy;

KW neuroprotective; antiparkinsonian; protein therapy; EST;

KX expressed sequence tag; gene; ss.

OS Homo sapiens.

PN W020022660-A2.

PD 21-MAR-2002.

PF 10-SEP-2001; 2001WO-US26015.

PR 11-SEP-2000; 2000US-0659671.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;

PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;

XX WPI; 2002-292408/33.

DR P-PSDB; ABB97448.

XX An isolated polynucleotide for treating diseases associated with its encoded polypeptide such as cancer and multiple sclerosis -

PS Claim 1; SEQ ID NO 272; 509pp; English.

XX The present invention provides the protein and coding sequences of 444 novel human proteins. These were isolated from expressed sequences tags (ESTs). They can be used to stimulate cell growth, to regulate haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth e.g. in burn treatment, to regulate the immune system e.g. to treat multiple sclerosis, to regulate activin or inhibin e.g. to treat infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke and cancer, to screen for drugs, to treat inflammatory conditions e.g. rheumatoid arthritis, and to treat nervous system disorders e.g. Parkinson's disease. The present sequence is a coding sequence of the invention.

SQ Sequence 6614 BP; 2094 A; 1420 C; 1827 G; 1273 T; 0 other;

Alignment Scores:

Pred. No.: 1-48e-214 Length: 6614
Score: 4340.50 Matches: 970
Percent Similarity: 69.13% Conservative: 190
Best Local Similarity: 57.81% Mismatches: 408
Query Match: 53.77% Indels: 110
DB: 24 Gaps: 33

US-09-902-432-4 (1-1596) x ABB59861 (1-6614)

QY 1 MetGlyAlaGlySerThrGluGlnArgSerProGluGln---ProAlaGlySerAsp 19
Db 192 ATGGCGCGCGGAGTCCACCGAGCAGCGCGGAGCAGCGCGCGGAGGAGTCC 251
QY 20 ThrProSerGluLeuValLeuSerGlyHisGlyProAlaGluAlaSer---GlyAla 38
Db 252 ACCCGCGCTGAGCCGAGCCAGCGCGCGCGCGCTCGCGGAGGCGCGCGAGCACC 311
QY 39 AlaGlyAspProAlaAspAla-----AspProAlaThrLysValLeuProGlnLysAsnGly 56
Db 312 ACCCGGACCCCGCCATCGCTCGGACCCCGCCACCCAGCTCTACAGAGAATGGT 371
QY 57 GlnLeuSerSerValAsnGlyValAlaGluGlnGlyAspValHisValGlnGluGluAsn 76
Db 372 CAGCTGCCACCATCAATGCGGTAGCTGAGCAAGATGAGCTCAGCCTCCAGGAGGTGAC 431
QY 77 GlnGluGlyGln-----GluGlu 82
Db 432 CTAATGGCCAGAAAGGAGGCCCTGAACGGTCAAGGAGCCCTAAACAGCCAGGAGGAAGA 491
QY 83 GluValValAspGluAspValGlyGlnArgGluSerGluAspValArgGluLysAspArg 102
Db 492 GAAGTCAATTGTCAGAGGTTGACAGAGAGACTCTGAAGATGTGAGCGGAAGAGACTCC 551

QY 103 ValGluGluMetAlaAlaAsnSerThrAlaValGluAspIleThrLysAspGlyGlnGlu 122
Db 552 GATAAAGAGATGGCTACTAAGTCAGCGGTGTTTCAGCAGATCAGCATGATGGCGAGGAG 611
QY 123 GluThrSerGluIleIleGluGlnIleProAlaSerGluAsnValGluGluMetVal 142
Db 612 GAGACACCCGAAATATCGAACAGATCTCTCTTCAGAAAGCAATTTAGAGAGCTAACA 671
QY 143 GlnProAlaGluSerGlnAlaAsnAspValGlyPheLysLysValPheLysPheValGly 162
Db 672 CAACCCACTGAGTCCCGAGCTAATGATTTTGAAGAGGTGTTTAAAGTTTGTGGC 731
QY 163 PheLysPheThrValLysLysAspLysAsnGluLysSerAspThrValGlnLeuLeuThr 182
Db 732 TTTAAATTCACCTGTGAAAAAGGATAAGACAGAGAACCTGACACTGTCAGCTACTCACT 791
QY 183 ValLysLysAspGluGlyGluGlyValAlaGluAlaSerValGlyValAlaGlyAspHisGlnGlu 202
Db 792 GTGAAGAAGATGAGGGAGGGAGCA-----GCAGGGCTGGCGACCAACAGGAC 842
QY 203 ProSerValGluThrAlaValGluGlyGluSerAlaSerLysGluSerGluLysGlnSer 222
Db 843 CCCAGCCTT-----GGGCTGGAGAAGCAGCATCCAAAGAAAGCGAACCCAAACAATCT 896
QY 223 ThrGluLysGlnGlyThrLeuLysGlnGluGlnSerThrGluLeuProLeuGln 242
Db 897 ACAGAGAAACCCGAGAGACCTCTGAAGCTGTAGCAAAAGCCACGCAAAAATTTCTCCCCA 956
QY 243 AlaGluSerAspGlnAlaAlaGluGluGluAlaLysAspGluGlyGluGlyGlnGlu 262
Db 957 GCCGAATCTGGCCAGCAGTG---GAGGAATGCAAGAGGAGGAGGAGAGAAACAGAA 1013
QY 263 LysGluProThrLysSerProGluSerProSerProValAsnSerGluThrThrSer 282
Db 1014 AAAGAACCTAGCAAGCTCAGCAATCTCCGACTAGTCCCGTACCACTGAAACAGGATCA 1073
QY 283 SerPheLysPhePheThrHisGlyThrAlaGlyThrArgLysLysThrSerPheLys 302
Db 1074 ACCTTCAAAAAATTTCTTCACTCAAGGTTGGCGCGCTGGCGCAAAAAGACCACTTCAGG 1133
QY 303 LysSerLysGluAspLeuGluThrAlaGluLysArgLysGluGlnGluAlaGluLys 322
Db 1134 AAGCCGAGGAGGATGAAGTGAAGCTTCAGAGAGAAAGAAAGGAGCAAGAGCAGAGAAA 1193
QY 323 ValAspGluGluGluLysGluLysThrGluProAlaSerGluGlu----- 337
Db 1194 GTAGACACAGAGAGACGCGAAAGGCGAGGTTGCTCCGAGAAACTGACCGCTCGAG 1253
QY 338 -----GlnGluProAlaGluAspThrAspGlnAlaArgLeuSerAlaAspTyr 353
Db 1254 CAAGCCCAACCCAGAGCGCGCAGAAAGTCCACAGAGCCCGGTTATCAGCTGAATAT 1313
QY 354 GluLysValGluLeuProLeuGluAspGlnValGlyAspLeuGluAlaSerSerGluGlu 373
Db 1314 GAGAAAGTTGAGCTGCGCTCAGAGGAGCAAGTCAAGTGGCTCGCAGGAGACCTTCTGAAG 1373
QY 374 LysCysAlaProLeuAlaThrGluValPheAspGluLysMetGluAlaHisGln---Glu 392
Db 1374 AAACCTGCTCCGTTGGCGACAGAGTGTGTTGATGAGAAATAGAGATCCACCAAGAGAG 1433
QY 393 ValValAlaGluValHisValSerThrValGluLysThrGluGluGlnGlyGlyGly 412
Db 1434 GTTGTGGCCGAAGTCCACGCTGAGCAGCGGTGGAGGAGAGAAACCGAAGCAGAG----- 1484
QY 413 GlyGluAlaGluGlyValValValGluGlyThrGlyGluSerLeuProGluLys 432
Db 1485 -----AAAAACGAGGTGGAGAGAAACACAGAGAGGCTCTGTCAGCTGAAGAA 1529
QY 433 LeuAlaGluProGlnGluValProGlnGluAlaGluProAlaGluGluMetLysSer 452
Db 1530 TTGGTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1589

QY 453 ArgLumecCysValSerGlyValAspHisThrGlnLeuThrAspLeuSerProGluGlu 472
 Db 1590 AAAGAAACGTTGTTCCGAGAGAGACCTACACAGAGAGCTGACTCATGTATGTAG 1649
 QY 473 LysThrLeuProLysHisProGluGlyLeuValSerGluValGlnLeuSerSerGln 492
 Db 1650 AAGGTGCTGTCCAAACCCCGGAGAGCGCTTGTCAGTGAAGTGGAAATGCTGTCACTACAG 1709
 QY 493 GluArgLysValGlnGlySerProLeuLysLeuPheSerSerSerGlyLeuLys 512
 Db 1710 GAGAGAAATGAAGTGCAGAGGAAGTCCACTAAAGAAAGCTTTTATCCAGCATCGCTTAAAA 1769
 QY 513 LysLeuSerGlyLysLeuSerGlnLysGlyValArgGlyGlyValAspGluGluProGly 532
 Db 1770 AAGCTTTCTGGAAAGAAAGAAAGGAAAGAA--GAGAGAGAGACAGAGAAATCAGGG 1826
 QY 533 GluTrpGlnHisLeuHisThrLeuSerProGluSerAlaAspGluGlnLysGlyLys 552
 Db 1827 GAGCACACTCAGGTTCCAGCCCATTTCTCCGACAGCCAGAGAGAGAAAGGCGAGAGC 1886
 QY 553 SerAlaSerSerProGluGluProGluGluThrCysLeuGluLysGlyProLeuGlu 572
 Db 1887 TCTGCTCATCCCTCGAGAGAGCCGAGAGATCAGTGTCTGAAAAGGGCTTAAAGCCGAG 1946
 QY 573 AlaProGlnAspGlyGluAlaGluGluGlyThrThrSerAspGlyGluLysLysArgGlu 592
 Db 1947 GTGACAGCAGAGTGGGAAAGCTGAAGAGAGCTACTTCGATGAGAGAAAGAAAGAGA 2006
 QY 593 GlyLeuThrProTrpAlaSerPheLysLysMetValThrProLysLysArgValArgArg 612
 Db 2007 GGTCATCCTCCCTGGGATCATTTCAAAAGATGTGTACCCCAAGAGCGTGTAGACGG 2066
 QY 613 ProSerGluSerAspLysGluGluGluLeuGlyValLysSerAlaThrLeuSerSer 632
 Db 2067 CCTTCGAAAGATGATTAAGAGATGAGCTGACAAGAGTCAAGAGCGCTTGTGTCTCC 2126
 QY 633 ThrAspSerThrValSerGluMetGlnAspGluValLysThrValGlyGluGluGlnLys 652
 Db 2127 ACCGAGAGCAGCGCTCTTAATGCAAGAAATGAAGAAAGGAGCGCTGAAAGCCAAAG 2186
 QY 653 ProGluGluProLysArgArgValAspThrSerValSerTrpGluLysLeuLeuLysCysVal 672
 Db 2187 CCGAGAGAACCAAGGCGCAAGGTGATCTCATGATCTTTGGAGACTTTAATTTGTG 2246
 QY 673 GlySerSerLysLysArgLysArgLysLysSerSerSerAspAspGluGlyProArg 692
 Db 2247 GGATCATCCAAAGAAAGAGCAAGAGAGAGTCTCTTGTGATGAGAAAGGGGACCAAAA 2306
 QY 693 ThrLeuGlyLysAspSerHisArgLysGluLysLysSerLysAspLysGluAlaGlyThr 712
 Db 2307 GCAATCGGAGAGACCAACAGAAAGCTGATGAGGCGGAAAGAAAGCAAGAGCGGGACA 2366
 QY 713 AspAlaValProAlaSerThrGlnGluGlnGluAlaGlnLysSerSerProGlu 732
 Db 2367 GACGGGATCTCTGCTGTTCCCAAGAACATGATCCAGGGCAGGGAAGTCTCTCCCGGAG 2426
 QY 733 ProAlaGlySerProSerGluGlyGluGlyValSerThrTrpGluSerPheLysArgLeu 752
 Db 2427 CAAGCTGGAAGCCCTACGAAAGGAGGCGTTTCCACCTGGGAGTCATTTAAAGGTTA 2486
 QY 753 ValThrProArgLysLysSerLysSerLysLeuGluGluLysAlaGluLysPhe 769
 Db 2487 GTACACCCCAAGAAACCAAAAGTCCAAAGCTGGAAGAGAAAGCAAGAACCTCCATAGCT 2546
 QY 770 ---SerSerValGluGlnLeuSerThrGluLeuGluProSerArgGluGluSerTrpVal 788
 Db 2547 GGGTCTGGTGAACATTTCCATCCAGACACTGAACCCGGTAAAGATCTCTGGGCTC 2606
 QY 789 SerLysLysPheLeuProGluArgArgLysLysArgAlaAspGlyLysGlnGluGln 808
 Db 2607 TCATATCAAAAGATTATTTCTGAGCAGAGAGAAAGGAGGAGATGGGAAACAAAGAACAA 2666
 QY 809 AlaThrValGluAspSerGlyProValGluLysGlnLysAspSerProAsnValProAla 828

Db 2667 GCCCTGTGTAAGAGCAGAGGCGCAACAGAGGCGCAACGAAGTACTGTGATGTCCGCGC 2726
 QY 829 ValValProLeuSerGlyTrpAsnAlaValGluArgGluLysMetGlu-----Ala 845
 Db 2727 GTGGTCCCTGTGTGATATGATGATCTGTAGAAAGGAGAAATGAGAGCAGCAAGCACC 2786
 QY 846 GlnGlyAsnThrGluLeuProGlnLeuGlnValAlaValSerGluGluLeuSer 865
 Db 2787 CAAAAGGCGCAGAGCAGCCCGAGAGAGGAGCAGCAGTGAAGTGTCCAGAGAGCTCAGC 2846
 QY 866 LysThrLeuValHisThrValSerValAlaValLysAspGlyThrArgAlaValThrSer 885
 Db 2847 GAGATCAGGTTCAATATATGACAGCAGCTGTCTGTACGAGAGAGAGGAGGAGGAGCTTCCATT 2906
 QY 886 ValGluGluArgSerProSerTrpLysSerAlaSerValThrGluProLeuGlnHisThr 905
 Db 2907 ATTGAAGAAAGGCTCTCTTGTGATATCTGTCTGATGAGAGAACTCTTGAAACAGTA 2966
 QY 906 AlaGlyGluAlaMetProProValGluGluValThrGlyLysAspLysLeuAla---Glu 924
 Db 2967 GAAGCTGAAGCCGACCTGTAACTGAGAGAGTATTGAAAGAAAGATTAATGCAAGAA 3026
 QY 925 GluThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHisAspAspMetVal 944
 Db 3027 GAACCCCCACGATTACTGAACCTTGCACAGAAACAGAGAGAGGCGGAGCAGGCTC 3086
 QY 945 ThrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluLysLeu 964
 Db 3087 GTTAGTGAGCGGAATTAACCCCGAAAGCTGTGACAGCTGCAGAAATCGACAGGCGCATTT 3146
 QY 965 ArgThrGluGluValThrGluAlaSerGlyValGluGluThrThrAspMetValSerAla 984
 Db 3147 GGTTCCGAAGAGAAAGCAAGAGCATCTGCTGTGAAGAGACACAGAAATGTGTACACA 3206
 QY 985 ValSerGlnLeuThrAspSerProAspThrThrGlnGluAlaThrProValGlnGluVal 1004
 Db 3207 GTTCCCGATTAAACGATCTCCACAGACACACAGAGAGGCGACTCCGTGACAGAGG 3266
 QY 1005 GluSerGlyValLeuAspThrGluGluGluGluArgGlnThrGlnAlaThrLeuGlnAla 1024
 Db 3267 GAAGGTGGCGTACGTGACATTAAGAAAGAGAGAGGCGGAGTCCAGAGTCTCTCCAGGCA 3326
 QY 1025 ValAlaAspLysValLysGluGluSerGlnValProAlaThr----- 1038
 Db 3327 GTGGCAGAAAGAGTAAAGAGAAATCCAGGCTGTGCACCGGTGCGCAGAAAGATG 3386
 QY 1039 ---GlnThrValGlnArgThrGlySerLysAlaLeuGluLysValGluGluValGluGlu 1057
 Db 3387 CTTCAGCTGTGCAGAGAGCA-----GAGGCAAGAAAGACCAAGAGAG 3428
 QY 1058 AspSerGluValLeuAlaSerGluLysGluLysAspValMetProLysGlyProValGln 1077
 Db 3429 CAGGCTGAAGCGTGGGTGTAAGAAAGAGACGATGATGTTGAAAGTATGATCTCAG 3488
 QY 1078 GluAlaGlyAlaGlnHisLeuAlaGlnGlySerGluThrGlyGlnAlaThrProGluSer 1097
 Db 3489 GAGGCAAAAACCTGACCTTTTACACAGAGGAGAGTGTGTGGGCGACACACCCCAAAAAGC 3548
 QY 1098 LeuGluLysValProGluValThrAlaAspValAspHis-----ValAlaThrCys 1113
 Db 3549 TTTGAAAAAGCTCTCAAGTGCACAGAGACATAGAGTCCAGTGAAGCTTTGATCCACTTGT 3608
 QY 1114 GlnVal-----LysLeuGlnGlnLeu---MetGluGlnAlaValAla 1127
 Db 3609 CAAGCGAAACCTTACGTGGGTAAATATCACAGAGATGATGAAACAGGCTATCCCC 3668
 QY 1128 SerGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrProLeuAlaAsp 1147
 Db 3669 CCTGACTCGGTGAAGACCTTACAGACAGTGAAGTGAAGAGACCCCGCTAGCCGAGC 3728
 QY 1148 SerAspThrAlaAspGlyThrGlnGlnAspGluThrLysAspSerGlnAspSerLysAla 1167

Db 3729 TTTGACGACACGACCAACCCAGAAAGACGAGATTGTGGAAATCCATGAGGAATGAG 3788
Qy 1168 ThrAlaValArgGlnSerGlnValThrGluGluGluAlaAlaThrAlaGlnLysGlu 1187
Db 3789 GTGCGATCTGTGATCCAGTCCAGGGGACAGAGACGAGGAGTCTCTGACAGAAAG 3848
Qy 1188 GluProSerThrLeuProAsnValProAlaGlnGluHisGlyGluGluProGly 1207
Db 3849 AGGCTCCAGCA---CCTTCAGTTTGTGTCAGGAAGAACTAAAGAAACAAATCAAG 3905
Qy 1208 ---ArgAspValLeuGluProThrGlnGlnGluLeuThrAlaAlaValProValLeu 1226
Db 3906 ATGGAAGACACTCTAGACATACAGATAAGAGGTGTGAGTGGAACTGTATCCATCTG 3965
Qy 1227 AlaLysThrGluValGlyGlnGluGlyGluValAspTrpLeuAspGlyGluLysValLys 1246
Db 3966 TCAAAGACTGAGGGAGCTCAA-----GAGGCTGACCATGTGTGATGAGAAACCAA 4019
Qy 1247 -----GluGluGlnGluValPheValHisSerGly-----ProAsnSer 1259
Db 4020 GACGTACCATTTTCGAAGAGCTTGAGGGGTCTATAGACACAGGCATPAACAGTCAGTCGG 4079
Qy 1260 GlnLysAlaAlaAspValThrTyAspSerGluValMetGlyValAlaGlyCysGlnGlu 1279
Db 4080 GAAAGGTCACTGAGTGGCTTAAAGGTGAAGGACAGAGAGCTGAATGTAAAGAG 4139
Qy 1280 LysGluSerThrGluValGlnSer-----LeuSerLeuGluGluGlyGlu 1294
Db 4140 GATGATGCTCTTGAATGACAGCTGACGCTAAGTCTCTCCATCCCTCCGCGGAGAGAG 4199
Qy 1295 MetGluThrAspValGluLysGluLysArgGluThrLysProGluGlnValSerGluGlu 1314
Db 4200 ATGCTAGTTCAAGTTCGAAGGAGAAACAGACAGACAGCAACCCATGTGAATGAAGAG 4259
Qy 1315 Gly---GluGlnGluThrAlaAlaProGluHisGluGlyThrTyGlyLysProValLeu 1333
Db 4260 AAGCTTGAGCAGCAAGACAGCTGTACGCTATCTGAAGAGTCAAGAGCTCCTCCAG 4319
Qy 1334 ThrLeuAspMetProSerSerGluArgGlyLysAlaLeuGlySerLeuGlyLys--- 1352
Db 4320 ACAGTGAATGTGCCCATCATAGATGGGCAAGAGAGTTCAGCAGTTTGGGAAGGAGCCT 4379
Qy 1353 ProSerLeuProAspGlnAspLysAlaGlyCysAlleGluValGlnValGlnSerLeuAsp 1372
Db 4380 CTTCCCTCCCTAGTCAAGAGGAGGAGTATGCACCAAAATTCAGATTTCAGAGCTCTGAG 4439
Qy 1373 ThrValThrGlnThrAlaGluAlaValGluLysValIle-----GluThrValVal 1390
Db 4440 GCATCATCTCTAAGCAGCGGTGACAGAGGAGAAAGGTCTTAGAGAAACTGCCAAC 4499
Qy 1391 IleSerGluThrGlyGluSerProGluCysValGlyAlaHisLeuLeuProAlaGluLys 1410
Db 4500 ATTTAGAAACAGGTGAACGTTGGAGCTGCAGTGCACATTTAGTTCTTGGGAAGAGAA 4559
Qy 1411 SerSerAlaThrGlyGlyHisTrpThrLeuGlnHisAlaGluAspThrValProLeuGly 1430
Db 4560 TCTCTGAAAAAATGAAGACTTTGCGCTCATCCAGGGGAAGATGTGTGCCACAGGG 4619
Qy 1431 ProGluSerGlnAlaGluSerIleProIleValThrProAlaProGluSerThrLeu 1450
Db 4620 CCCAGTGTGAGGAAATCGACACCATGTATGATCTCTACTACCAAGAAAGGCTTA 4679
Qy 1451 HisProAspLeuGlnGlyGluIleSerAlaSerGlnArgGluSerGluGluGluAsp 1470
Db 4680 AGTTCGACCTGGAGAGAGAGAAACCATCATCTAGTGGAGTGGAGTCAAGTCAATCGAT 4739
Qy 1471 LysProAspAlaGlyProAspAlaAspGlyLysGluSerThrAlaLeuGluValLeu 1490
Db 4740 GAGCAGGTTCTGTCAGGAGGAGTC-----AAAGTGAGTGTAGCAATTTAGGAGGATTA 4793
Qy 1491 LysAlaGluProGluIleLeuGluLysSerLysSerAsnLysIleValLeuAsnVal 1510
Db 4794 GAGCCTGAAATGGGATTTTGGAACTTGAGACCAAGCAAGCAATTAATGTGTCCAAACATC 4853

Qy 1511 IleGlnThrAlaValAspGlnPheAlaArgThr---GluThrAlaProGluThrHisAla 1529
Db 4854 ATCCAGACAGCCGTTGACAGGTTGTACGTACAGAAAGAAACACCCAGCAATTTGTGACG 4913
Qy 1530 TyrAspSerGlnThrGlnValProAlaCysArgLeuAspSerArgGluProAsnArgCys 1549
Db 4914 TCTGAGTTTACAGACACACAGCTACGTGATAAAAGCTGACAGCCAGGCGCTGGACAGGAA 4973
Qy 1550 TrpThrLys-----MetLysAspAla-LysMetLysHisProValProGlu 1564
Db 4974 ACGGAAAGAGAGAGAGGAGCCCTCAGGCTCTGCACAGATGAACACCAATTAATTCTCA 5033
Qy 1564 nProArgGlu-----AspLeuGlnValLeuThrValLeuGluAl 1577
Db 5034 GCCAAAGAGGAGTGCAGATCAACCGCAGTGGGACAAAGCACATTTCTGATATTTCCAAAGAC 5093
Qy 1577 aTrpAlaGlnProArgLysCysLeuProArgLeuGlnLeuLysAlaProVal 1594
Db 5094 ATGAGTGAAGCCCTCAGAAAGACCATGATGTTGAGGTAGAAGGTTCCACTG 5145
RESULT 4
ID ABN96942
XX ID ABN96942 standard; DNA; 6608 BP.
AC ABN96942;
XX
DT 13-AUG-2002 (first entry)
XX
DE Gene #3440 used to diagnose liver cancer.
XX
KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
metastatic liver tumour; cytostatic; expression profile; disease state;
disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
OS Homo sapiens.
XX
PN WO200229103-A2.
XX
PD 11-APR-2002.
XX
PF 02-OCT-2001; 2001WO-US30589.
XX
PR 02-OCT-2000; 2000US-237054P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
XX WPI; 2002-426119/45.
DR
XX
PT Diagnosing and detecting the progression of liver cancer,
hepatocellular carcinoma or metastatic liver tumor in a patient,
involves detecting the level of expression of two or more genes in a
liver tissue sample
XX
PS Claim 1; SEQ ID NO 3440; 298pp; English.
XX
CC The invention relates to a novel method for diagnosing and detecting the
progression of liver cancer, hepatocellular carcinoma or metastatic liver
tumor in a patient, and differentiating metastatic liver cancer from
hepatocellular carcinoma in a patient, involving detecting the level of
expression of two or more genes represented in ABN93503-ABN97455 in a
tissue sample. The method of the invention has hepatotropic, and
cytostatic activity. The method is useful for diagnosing and detecting
the progression of liver cancer, hepatocellular carcinoma and metastatic
liver carcinoma in a patient. The method is useful for identifying
expression profiles which serve as useful diagnostic markers as well as
markers that can be used to monitor disease states, disease progression,
drug toxicity, drug efficacy and drug metabolism.
CC Note: the sequence data for this patent did not form part of the printed
specification, but was obtained in electronic format directly from WIPO
at ftp.wipo.int/pub/published_pct_sequences.

Db 2124 ACCGAGACAGCCTCTGAATGCAAGAAATGAAGGAGCGTGGAGAGCCAAAG 2183
QY 653 ProGluGluProLysArgArgValAspThrSerValSerTyrGluAlaLeuLeuLeuVal 672
Db 2184 CCGGAGAACCAAGCCAGGTCGATACCTCAGTATCTTGGGAAGCTTTAAATTTGTGTG 2243
QY 673 GlySerSerLysLysArgAlaArgLysAlaSerSerSerSerSerSerSerSerSer 692
Db 2244 GGAATCATCCAAAGAAAGAGCAAGGAGGAGGTCCTCTCTCTGATGAGGAGGAGGACCAAAA 2303
QY 693 ThrLeuGlyGlyAspSerHisArgAlaGluAlaSerLysLysLysLysLysLysLys 712
Db 2304 GCAATGGAGAGACACCAAGAAAGCTGTAGAGCCCGGAAAGACAAAGAGACGGGACCA 2363
QY 713 AspAlaValProLaserThrGlnGlnAspGlnAlaGlnGlnGlnGlnGlnGlnGln 732
Db 2364 GACGGGATCTTGCTGTTCCCAAGAACATGATCCAGGGCAGGGAAGTTCTCTCCCGGAG 2423
QY 733 ProAlaGlySerProSerGluGlyGluGlyValSerThrTyrGluSerPheLysArgLeu 752
Db 2424 CAAGCTGGAAGCCCTACCGAAGGGGAGGGGCTTTCCACCTGGGAGTCATTTAAAGGTTA 2483
QY 753 ValThrProArgLysLysSerLysSerLysLysLysLysLysLysLysLysLysLys 769
Db 2484 GTACGCGCAAGAAAAAATCAAGTCCAGCTGGAAGAAAAAGCGAAGACTCCATAGCT 2543
QY 770 ---SerSerValGluGlnLeuSerThrGluLeuProSerArgGluGluSerTyrVal 788
Db 2544 GGGTCTGGTGTAGAACATTCCTCCAGACACTGAACCCCGTAAAGAAAGATCTCGGGTC 2603
QY 789 SerIleLysLysPheIleProGlyArgArgLysLysArgAlaAspGlyLysGlnGln 808
Db 2604 TCAATCAAGAAGTTTATCTCGACGAAGCAAGAAAAAGGCCAGATGGCAACAAGAACAA 2663
QY 809 AlaThrValGluAspSerGlyProValGluIleAsnGluAspAspProAsnValProAla 828
Db 2664 GCCCCTGTTGAAGACGCGAGGCCCAACAGGGGCCAACGAAGATGACTCTGATGTCCCGGCC 2723
QY 829 ValValProLeuSerGluTyrAsnAlaValGluArgGluLysMetGlu-----Ala 845
Db 2724 GTGGTCCCTCTGCTGATGATGCTGTAGAAAGGAGAAATGGAGGCACAGCAAGCC 2783
QY 846 GlnGlyAsnThrGluLeuProGlnLeuLeuGlyAlaValTyrValSerGluGluLeuSer 865
Db 2784 CAAAAAGGCGCAGACGAGCCGAGCAGACAGAGCCACTCAGGTGTCCAAGAGCTCAGC 2843
QY 866 LysThrLeuValHisThrValSerValAlaValIleAspGlyThrArgAlaValThrSer 885
Db 2844 GAGAGTCAGGTTTCATATGATGGCAGCAGCTGCTGACCGGACGAGGCGAGCTTACCATT 2903
QY 886 ValGluGluArgSerProSerTyrIleSerAlaSerValThrGluProLeuGluHisThr 905
Db 2904 ATTGAAGAAAGGTCCTCTCTTGATATCTGCTTCACTGACAGAACCTCTTGAACAAGTA 2963
QY 906 AlaGlyGluAlaMetProProValGluGluValThrGluLysAspIleAla---Glu 924
Db 2964 GAAGCTGAAGCCGACCTGTTAACTGAGGAGGTATTGGAAGAGAAAGTAATTGCAGAAAGAA 3023
QY 925 GluThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHisAspAspMetVal 944
Db 3024 GAACCCCCACCGTTACTGAACCTCTCCACAGAACACAGAGAGCGCCGGGCGCACCGGTC 3083
QY 945 ThrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeu 964
Db 3084 GTTAGTGAGGCGGAATTTGACCCCGAAGCTGTGACAGCTGCAGAAATCGCAGGCGCCATTG 3143
QY 965 ArgThrGluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSerAla 984
Db 3144 GGTTCGGAAGAGAACCGAAGCATCTGCTGCTGAAGAGACACAGAAATGGTGTACAGCA 3203
QY 985 ValSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGluVal 1004

Db 3204 GTCTCCAGTTAAACGACTCCCCAGACACCAACAGAGAGGCCACTCGGTGCGAGAGGTG 3263
QY 1005 GluSerGlyValLeuAspThrGluGluGluArgGlnThrGlnAlaIleLeuGlnAla 1024
Db 3264 GAAGGTGGCGTACCTCAGATAGAAAGAGAGAGAGCGGACTCAAGAGGTCCTCCAGGCA 3323
QY 1025 ValAlaAspLysValLysGluGluSerGlnValProAlaThr----- 1038
Db 3324 GTGGCAGAAAAGTGAAGAGGAATCCAGCTCGCTGGCACCGGTGGGCCAGAGATGTG 3383
QY 1039 ---GlnThrValGlnArgThrGlySerLysAlaLeuGluLysValGluGluValGluGlu 1057
Db 3384 CTTTACGCTGTGCAGAGCA-----GAGGCAGAAAGACACAGAAAG 3425
QY 1058 AspSerGluValLeuAlaSerGluLysGluLysValMetProLysGlyProValGln 1077
Db 3426 CAGGCTGAAGCGTCGGTCTGAAGAAAGAGAGCGGATGTAGTGTGAAGTAGATGCTCAG 3485
QY 1078 GluAlaGlyAlaGluHisLeuAlaGlnGlySerGluThrGlyGlnAlaThrProGluSer 1097
Db 3486 GAGGCAAAAACCTGAGCCTTTTACACAAAGGGAAGGTGGTGGGGCAGACCACCCAGAAAGC 3545
QY 1098 LeuGlu---ValProGluValThrAlaAspValAspHis-----ValAlaThrCys 1113
Db 3546 TTTGAAAAGACTCTCAAGTCACAGAGCATAGAGTCCAGTGAGCTTTGTAACCATTTGT 3605
QY 1114 GlnVal-----IleLysLeuGlnGlnLeu---MetGluGlnAlaValAla 1127
Db 3606 CAAGCCGAAACCTTAGCTGGGTAAATACAGAGAGATGGTGTGAACAGGCTATCCCC 3665
QY 1128 ProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrProLeuAlaAsp 1147
Db 3666 CTTGACTCGGTGAAAAACCTCAGACAGTGTGAGTGTGGAAGCACCCTCGTAGCCGAC 3725
QY 1148 SerAspThrAlaAspGlyThrGlnGlnAspGluThrIleAspSerGlnAspSerLysAla 1167
Db 3726 TTTGAGCGCACCGGCAACCCAGAAAGAGAGATTTGGAAATTCATGAGGAAATGAG 3785
QY 1168 ThrAlaAlaValArgGlnSerGlnValThrGluGluGluAlaAlaThrAlaGlnLysGlu 1187
Db 3786 GTCGCATCTGGTATCCAGTCAGGGGCAACAGAGCAGAGAGGAGGCTTCTGCACAGAAAGAG 3845
QY 1188 GluProSerThrLeuProAsnAsnValProAlaGlnGluHisGlyGluGluProGly 1207
Db 3846 AGGCCTCCAGCA---CCTTCAGTTTGTTCAGGAAGAAACTAAAGAAACAATCAAAG 3902
QY 1208 ---ArgAspValLeuGluProThrGlnGlnGlnLeuThrAlaAlaValProValLeu 1226
Db 3903 ATGGAAAGACACTCTAGAGCATACAGATAAAGAGGTGTCTAGTGGAAACTGTATCCATTCTG 3962
QY 1227 AlaLysThrGluValGlyGlnGluGlyGluValAspTyrLeuAspGlyGluLysValLys 1246
Db 3963 TCAAGACTGAGGGGACTCAA-----GAGGCTGACCATGATCTGATGAGAAACCAAA 4016
QY 1247 -----GluGluGlnGluValPheValHisSerGly-----ProAsnSer 1259
Db 4017 GACGTACCATTTTTTCAAGAGCATTTGAGGGTCTATAGACACAGGCATACACAGTCAGTCGG 4076
QY 1260 GlnLysAlaAlaAspValThrTyrAspSerGluValMetGlyValAlaGlyCysGlnGlu 1279
Db 4077 GAAAAAGGTCACTGAAGTTCCTTTAAAGGTGAAGGGGACAGAAAGAGCTGAATGTAAAGAG 4136
QY 1280 LysGluSerThrGluValGlnSer-----LeuSerLeuGluGluGlyGlu 1294
Db 4137 GATGATGCTCTTGAATGACAGTACAGTACGTTAAGTCTCTCCATCCCTCCCTGGGAGAGAGAG 4196
QY 1295 MetGluThrAspValGluLysGluLysArgGluThrLysProGluGlnValSerGluGlu 1314
Db 4197 ATGTTAGTTCAAGTCCGAAGGGAGAAAAACAGAGCAGAGCCACCCATCTGTAATGAAGAG 4256
QY 1315 Gly---GluGlnGluThrAlaAlaProGluHisGluGlyThrTyrGlyLysProValLeu 1333
Db 4257 AAGCTTGAGCAGCAACAGCTGTTACCGTATCTGAAAGAGGTGTAAGAGGCTCTCTCCAG 4316


```

QY 1334 ThrLeuAspMetProSerSerGluArgGlyLysAlaLeuGlySerLeuGlySer--- 1352
Db 4317 ACAGTGAATGTCCCATCATCATGATGAGGCAAGAGAAAGAGTTCAGTTCAGGAAAGCCCT 4376
QY 1353 ProSerLeuProAspGluAspLysAlaGlyCysIleGluValGluValGlnSerLeuAsp 1372
Db 4377 CCTCCCTGCTAGTGTCAAGAGGAGGATGATATGACCAAAATTCAGTTCAGGCTCTGAG 4436
QY 1373 ThrThrValThrGlnThrAlaGluValAlaValGluLysValIle-----GluThrValVal 1390
Db 4437 GCATATTACACTCTAAACAGCGGCTGACAGAGAGAAAGGCTCTTAGAGAAATCTGCACAA 4496
QY 1391 IleSerGluThrGlyGluSerProGluCysValGlyAlaHisLeuLeuProAlaGluLys 1410
Db 4497 ATTTAGAAACAGGTGAACCTTTGGAGCCTGACAGGTCACATTTAGTTCGAAAGGAAA 4556
QY 1411 SerSerAlaThrGlyGlyHisThrThrLeuGlnHisAlaGluAspThrValProLeuGly 1430
Db 4557 TCCCTGAAAAAATGAAGACTTTGCGCTCATCCAGGGAAAGATGCTGTGCCACAGGG 4616
QY 1431 ProGluSerAlaAlaGluSerIleProIleIleValThrProAlaProGluSerThrLeu 1450
Db 4617 CCCGACTGTACAGGCAAAATCGACACCGATGATATCTGCTACTCCAAAGAAAGCTTA 4676
QY 1451 HisProAspLeuGlnGlyGluIleSerAlaSerGlnArgGluArgSerGluGluAsp 1470
Db 4677 AGTCCGACCTGGAAGAGAGAAACACATCATCTGAATGGAAGTCAGATGAATCGAT 4736
QY 1471 LysProAspAlaGlyProAspAlaAspGlyLysGluSerThrAlaIleGluLysValLeu 1490
Db 4737 GAGCAGGTTGCTGCCAGAGGTC-----AAAGTAGGTGACAAATTGAGATTTA--- 4787
QY 1491 LysAlaGluProGlu-----IleLeuGluLeuGluSerLysSerAsnLysIleValLeu 1508
Db 4788 -----GAGCCGAAAAATGGATTTTGGAACTTGGAAACCAAAAGCACTTAATGTCCAA 4841
QY 1509 AsnValIleGlnThrAlaValAspGlnPheAlaArgThr--GluThrAlaProGluThr 1527
Db 4842 AACATCATTCACAGACCCCTTGACCTGATTCAGTACGAAAGAAACACCCAGCCGAAAG 4901
QY 1528 HisAlaThrAspSerGlnThrGlnValProAlaCysArgLeuAspSerArgGluProAsn 1547
Db 4902 TTGACCTGTGAAGTTACGACACAAAGCTCAGTGATAAAGCTGACGCAAGCGCTGGA 4961
QY 1548 ArgCysTrpThrLys-----MetLysAspAla-LysMetLysHisProVa 1562
Db 4962 CAGGAAACGAGAAAGAGAGAGAGAAAGAACTCAGGCTCTGCACAGATGAACACCAATT 5021
QY 1562 IProGlnProArgGlu-----AspLeuGlnValLeuThrValLe 1575
Db 5022 ACTTCAGCCAAAGAGAGTACAGATCAACCGACGAGTGCAAGACATCTCTGATTATTC 5081
QY 1575 uGluAlaTrpAlaGlnProArgLysCysLeuProArgLeuGlnLeuLysAlaProVal 1594
Db 5082 AAAGACATGATGAGCTTCAGAAAGACATGACTGTGAGTAGAGGTTCCACTG 5139

```

```

XX 00200194629-A2.
XX 13-DEC-2001.
XX 30-MAY-2001; 2001WO-US10838.
XX 05-JUN-2000; 2000US-208473P.
XX 05-JUN-2000; 2000US-209531P.
XX 18-SEP-2000; 2000US-233133P.
XX 18-SEP-2000; 2000US-233617P.
XX 20-SEP-2000; 2000US-234009P.
XX 20-SEP-2000; 2000US-234034P.
XX 20-SEP-2000; 2000US-234052P.
XX 22-SEP-2000; 2000US-234509P.
XX 22-SEP-2000; 2000US-234567P.
XX 25-SEP-2000; 2000US-234923P.
XX 25-SEP-2000; 2000US-234924P.
XX 25-SEP-2000; 2000US-235077P.
XX 25-SEP-2000; 2000US-235082P.
XX 25-SEP-2000; 2000US-235134P.
XX 25-SEP-2000; 2000US-235280P.
XX 26-SEP-2000; 2000US-235637P.
XX 26-SEP-2000; 2000US-235638P.
XX 27-SEP-2000; 2000US-235711P.
XX 27-SEP-2000; 2000US-235720P.
XX 27-SEP-2000; 2000US-235840P.
XX 27-SEP-2000; 2000US-235863P.
XX 28-SEP-2000; 2000US-236028P.
XX 28-SEP-2000; 2000US-236032P.
XX 28-SEP-2000; 2000US-236033P.
XX 28-SEP-2000; 2000US-236034P.
XX 28-SEP-2000; 2000US-236109P.
XX 28-SEP-2000; 2000US-236111P.
XX 29-SEP-2000; 2000US-236842P.
XX 29-SEP-2000; 2000US-236891P.
XX 02-OCT-2000; 2000US-237172P.
XX 02-OCT-2000; 2000US-237173P.
XX 02-OCT-2000; 2000US-237278P.
XX 02-OCT-2000; 2000US-237294P.
XX 02-OCT-2000; 2000US-237295P.
XX 02-OCT-2000; 2000US-237316P.
XX 03-OCT-2000; 2000US-237425P.
XX 03-OCT-2000; 2000US-237598P.
XX 03-OCT-2000; 2000US-237604P.
XX 03-OCT-2000; 2000US-237606P.
XX 03-OCT-2000; 2000US-237608P.
XX 01-NOV-2000; 2000US-244867P.
XX 01-NOV-2000; 2000US-245084P.
XX (AVAL-) AVALON PHARM.
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
XX Soppet DR, Weaver Z;
XX WPI; 2002-186264/24.
XX
XX Screening for anti-neoplastic agent involves exposing cells to a
XX chemical agent to be tested for anti-neoplastic activity, and
XX determining a change in expression of a gene of a signature gene set -
XX Claim 1; SEQ ID 275; 44p; English.
XX
XX The present invention describes a method (M1) for screening for an
XX anti-neoplastic agent. The method involves exposing cells to a chemical
XX agent to be tested for anti-neoplastic activity, determining a change in
XX expression of at least one gene (I) of a signature gene set, where (I)
XX comprises a sequence (S) selected from 8447 sequences (given in ABL61664
XX to ABL70110), or is at least 95% identical to (S), where a change in
XX expression is indicative of anti-neoplastic activity. (I) has cytostatic
XX activity and can be used in gene therapy. M1 can be used for screening
XX an anti-neoplastic agent, and can be used for producing a product which
XX is the data collected with respect to the anti-neoplastic agent as a

```


QY 1295 MetGluThrAspValGluLysGluArgGluThrLysProGluGlnValSerGluGlu 1314
 DB 4197 ATGTAGTTCAGTCGAAGAGGAGAAACAGACGACGACCAACCCATGTAATGAAGAG 4256
 QY 1315 Gly---GluGlnGluThrAlaAlaProGluHisGluGlyThrTyrglyLysProValLeu 1333
 DB 4257 AAGCTTGAGCAGCAACAGCTGTTACCGTATCTGAAGAGGTCTAGTAGAGCTCTCCAG 4316
 QY 1334 ThrLeuAspMetProSerSerGluArgGlyLysAlaLeuGlySerLeuGlyGlySer--- 1352
 DB 4317 ACAGTGAATGTGCCCATCATAGATGGGCAAGAGGAGTCTGAGAGGAGGAGGAGGAGG 4376
 QY 1353 ProSerLeuProAspGlnAspLysAlaGlyCysIleGluValGlnValGlnSerLeuAsp 1372
 DB 4377 CTCCCTCCCTAGTCAAGAGGAGGAGTATGACCAAAAAATCAAGTTTCAGAGCTCTGAG 4436
 QY 1373 ThrThrValThrGlnThrAlaGluAlaValGluLysValIle-----GluThrValVal 1390
 DB 4437 GCATCATTCCTCTACAGCGGCTGCAGAGGAGGAAAGGTCTTAGGAGAAACTGCCAAC 4496
 QY 1391 IleSerGluThrGlyGluSerProGluCysValGlyAlaHisLeuLeuProAlaGluLys 1410
 DB 4497 ATTTAGAAACAGGTGAACGTTGGAGCCTGCAGTGCACATTAGTTCTTGGAGAGGAAA 4556
 QY 1411 SerSerAlaThrGlyGlyHisTrpThrLeuGlnHisAlaGluAspThrValProLeuGly 1430
 DB 4557 TCCTCTGAAAAAATGAAGACTTTGCCGCTCATCCAGGGGAAGATGCTGTGCCACAGGG 4616
 QY 1431 ProGluSerGlnAlaGluSerIleProIleIleValThrProAlaProGluSerThrLeu 1450
 DB 4617 CCGAGCTGTGAGCAAAATGCACACAGTGTATCTCTACTACTACCAAGAAAGGCTTA 4676
 QY 1451 HisProAspLeuGlnGlyIleSerAlaSerGlnArgGluArgSerGluGluGluAsp 1470
 DB 4677 AGTTCGACCTGGAAGAGAGAGAAACACATCATCTGAAGTGAAGTGCAGATGAATCGAT 4736
 QY 1471 LysProAspLaglyProAspAlaAspGlyLysGluSerThrAlaIleGluLysValLeu 1490
 DB 4737 GAGCAGGTTCCTGCCAGGAGGTC-----AAAGTGAGTGTAGCAATTTGAGGATTTA-- 4787
 QY 1491 LysAlaGluProGlu-----IleLeuGluLeuGluSerLysSerAenLysIleValLeu 1508
 DB 4788 -----GAGCTGAAATGGGATTTTGGAACTTGAGACCAAAAGCAGTAACCTTGTCAC 4841
 QY 1509 AsnValIleGlnThrAlaValAspGlnPheAlaArgThr---GluThrAlaProGluThr 1527
 DB 4842 AACATCATCCAGACAGCGTGTACCACTTTGTACGACAGAGAAACAGCCACCGAAATG 4901
 QY 1528 HisAlaTyrrAspSerGlnThrGlnValProAlaCysArgLeuAspSerArgGluProAsn 1547
 DB 4902 TTGACGCTCTGAGTTACAGACACAAAGCTCACGTGATAAAGCTGACAGCCAGACGCTGGA 4961
 QY 1548 ArgCysTrpThrLys-----MetLysAspAla-LysMetLysHisProVa 1562
 DB 4962 CAGAAACGAGAGAAAGAGAGAGAAACCTCAGCGCTCTGCACAGGATGAAACACCAATT 5021
 QY 1562 lProGlnProArgGlu-----AspLeuGlnValLeuThrValLe 1575
 DB 5022 ACTTCAGCCAAAGAGGAGTACAGTCAACCGCAGTGGGACAGCACATCTTGATATTCC 5081
 QY 1575 uGluAlaTrpAlaGlnProArgLysCysLeuProArgLeuGlnLeuLysAlaProVal 1594
 DB 5082 AAAGACATGATGTAAGCCTCAGAAAAGACCATGACTGTTGAGGTAGAGGTTCCACTG 5139
 RESULT 6
 AAV23545
 ID AAV23545 standard; cDNA; 6605 BP.
 XX
 AC AAV23545;
 XX
 DT 13-JUL-1998 (first entry)
 XX
 DE Human gravin polypeptide coding sequence.

XX Gravin; kinase anchoring protein; type II regulatory subunit; PKA; PKC;
 KW cAMP-dependent protein kinase; protein kinase C; autoimmune disease;
 KW Myasthenia gravis; nicotinic acetylcholine receptor; ss.
 XX Homo sapiens.
 OS
 FH Key Location/Qualifiers
 FT CDS 192..5534
 FT /*tag= a
 XX
 PN US5741890-A.
 XX
 PD 21-APR-1998.
 XX
 PF 19-DEC-1996; 96US-0769309.
 XX
 PR 19-DEC-1996; 96US-0769309.
 XX
 PA (UYOR-) UNIV OREGON HEALTH SCI.
 PI Klauck TM, Nauert JB, Scott JD;
 XX
 XX WPI: 1998-260552/23.
 DR P-PSDB; AAW53863.
 XX
 PT New polypeptide fragments of protein kinase binding protein gravin -
 PT are useful for the study of modulation of action between gravin and
 PT protein kinase(s)
 XX
 PS Example 1; Column 19-32; 32pp; English.
 XX
 CC This sequence encodes the human gravin polypeptide, which represents
 CC a polypeptide of the invention. The polypeptides are fragments capable of
 CC binding to type II regulatory subunit of cAMP-dependent protein kinase
 CC (PKA). Gravin is a kinase anchoring protein that binds to type II
 CC regulatory subunits of PKA and protein kinase C (PKC). Gravin is also an
 CC antigen of the autoimmune disease Myasthenia gravis (MG), where a patient
 CC develops antibodies against their own nicotinic-acetylcholine receptors.
 CC The polypeptides are useful for providing analogues of gravin in the
 CC study of the modulation (e.g. blocking, inhibiting and stimulating) of
 CC interactions between gravin and kinase. The peptides are involved in the
 CC modulation of gravin-kinase interactions.
 XX
 SQ Sequence 6605 BP; 2091 A; 1416 C; 1826 G; 1272 T; 0 other;
 Alignment Scores:
 Pred. No.: 1.72e-211 Length: 6605
 Score: 4281.00 Matches: 965
 Percent Similarity: 68.63% Conservative: 189
 Best Local Similarity: 57.44% Mismatches: 409
 Query Match: 53.03% Indels: 117
 DB: 19 Gaps: 37
 US-09-902-432-4 (1-1596) x AAV23545 (1-6605)
 QY 1 MetGlyAlaGlySerThrGluGlnArgSerProGluGln---ProAlaGlySerAsp 19
 DB 192 ATGGGCGCGGGAGTCTCCACCGAGCAGCGAGCGCGCGGAGGAGGAGCTCC 251
 QY 20 ThrProSerGluLeuValLeuSerGlyHisGlyProAlaAlaGluAlaSer---GlyAla 38
 DB 252 ACGCCGGCTGAGCGCGAGCCAGCGCGCGCGCGCTCGCGCGGCGCGGCGGCGGAGCACC 311
 QY 39 AlaGlyAspProAlaAspAla-----AspProAlaThrLysLeuProGlnLysAsnGly 56
 DB 312 ACCGCGGAGCCCGGCCCATCTGCTCGGACCCCGCCACCAAGCTCTACAGAAATGGT 371
 QY 57 GlnLeuSerSerValAsnGlyValAlaGluGlnGlyAspValHisValGlnGluGluAsn 76
 DB 372 CAGCTGTCCACCATCATCGCTAGCTAGCAAGATGAGCTCAGCTCAGCGGCGGTGAC 431
 QY 77 GlnGluGlyGln-----GluGlu 82

```

Db 432 CTAATGCGCAGAAAGAGCCCTGAA CGCTCAAGAGCCCTTAAACAGCCAGAGAGAGAA 491
Oy 83 GluValAlaAspGluAspValGlyIleArgGluSerGluAspValArgGluIleuMetLysSer 102
Db 492 GAAAGCATTTGTCACGAGGTGGACAGAGAACTCTGAAAGATGTGAGCAAGAAAGAGCTCC 551
Oy 103 ValGluIleuMetAlaAlaAsnSerThrAlaValGluAspIleThrLysAspGlyGlyGlu 122
Db 552 GATAAAGAGATGGCTAATGATGAGCGGTGTGTTCACGATCAACATATGATGGCGAGAG 611
Oy 123 GluThrSerGluIleIleGluGluIleProAlaSerGluAsnAsnValGluIleuMetVal 142
Db 612 GAGAACCGGAAT--ATCCAAAGATTCCTTCTTCAGAAAGCAATTTAAGAGGCTAACA 668
Oy 143 GluProAlaGluSerGlnAlaAsnAspValGlyPheLysLysValPheLysPheValGly 162
Db 669 CAACCCACTGAGTCCAGGCTAATGATTTAAGAGGTGTTTAAAGTTGTTGCG 728
Oy 163 PheLysPheThrValLysLysAspLysAsnGluLysSerAspThrValGluIleuLeuThr 182
Db 729 TTTAAATTCATCTGTAAGAAAGATAGACAGAGAACCTGACACTGTCCAGCTACTCAGCT 788
Oy 183 ValLysLysAspGluGlyGlyGluGluIleGluIleAspValGlyAlaGluIleuMetVal 202
Db 789 GTGAAGAAAGATGAAGGAGGAGGAGCA-----CGAGGGGCTGGCGACCAACAGAGAC 839
Oy 203 ProSerValGluThrAlaValGlyGluSerAlaSerLysGluSerGluIleuLysGlnSer 222
Db 840 CCCAGCTT-----GGGGCTGGAGAGACGACATCCAAAGAAAGCAACCCAAACATCT 893
Oy 223 ThrGluLysGluGlyGlyThrLysLysGluGluGlnSerSerThrGluIleProLeuGln 242
Db 894 ACAGAGAAACCCGAGAGAGACCTGAAAGCTGAGCAAGCCAGCACGCAAAATTTCTCCCA 953
Oy 243 AlaGluSerAspGlnAlaIleGluGluGluAlaLysAspGluGlyGlyGluIleuLysGln 262
Db 954 GCGGAATTCGGCCAGACAGCTG--GAGGAATGCAGAAAGAGAGAGAGAGAGAGAGAGAA 1010
Oy 263 LysGluProThrLysSerProGluSerProSerSerProValAsnSerGluThrThrSer 282
Db 1011 AAGAACTTACAGCAAGTCTCCAGATCTCCGACTAGTCCGTGACGATGAAGAGATCA 1070
Oy 283 SerPheLysLysPhePheThrHisGlyTTPAlaGlyTTPArgLysLysThrSerPheLys 302
Db 1071 ACCTTCAAAAATTTCTTCACTCAAGGTTGGCGCGCTGGCGCAAAAAGACAGATTTCAAG 1130
Oy 303 LysSerLysGluAspAspLeuGluThrAlaGluLysArgLysGluGluGluIleuLys 322
Db 1131 AAGCCGAGAGAGATAGAGGAAAGCTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAA 1190
Oy 323 ValAspGluGluGluLysGluLysThrGluProAlaSerGluGlu-----337
Db 1191 GTAGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1250
Oy 338 -----GlnGluProAlaGluAspThrAspGlnAlaArgLeuSerAlaAspTyr 353
Db 1251 CAAGCCACCCACAGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1310
Oy 354 GluLysValGluIleuProLeuGluAspGlnValGlyAspLeuGluIleuSerSerGlu 373
Db 1311 GAGAAAGTTGAGTGCCTCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1370
Oy 374 LysCysAlaProLeuAlaThrGluValPheAspGluLysMetGluAlaHisGln--Glu 392
Db 1371 AAACCTGCTCCGTGGCGACAGAGAGTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1430
Oy 393 ValValAlaGluValHisValSerThrValGluLysThrGluGluGluGluGlyGly 412
Db 1431 GTTGTGGCGAGAGTCCACGTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1481
Oy 413 GlyGluAlaGluGlyGlyValValValGluGlyThrGlyGluSerLeuProProGluLys 432

```

```

Db 1482 -----AAACGAGGTGAGAGAAACAGAGAGGTCTGTGCCAGCTGAAGAA 1526
Oy 433 LeuAlaGluProGluGluValProGluGluAlaGluProAlaGluGluIleuMetLysSer 452
Db 1527 TTGTTGGAATGATGTCAGAACTCTCAGAGAGCCGAACCTGCCAAGAGAGCTGTGAAGCTC 1586
Oy 453 ArgGluMetCysValSerGlyGlyAspHisThrGluIleuThrAspLeuSerProGluGlu 472
Db 1587 AAGAAACGTGTGTTTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1646
Oy 473 LysThrLeuProLysHisProGluGlyLysValSerGluValGluIleuMetLysSerGln 492
Db 1647 AAGGTCTGTCCAAACCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1706
Oy 493 GluArgLysValGluGlySerProLeuLysLysLeuPheSerSerSerGlyLysLys 512
Db 1707 GAGAGAAAGAGAGTGCAGAGAGAGTCACTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1766
Oy 513 LysLeuSerGlyLysLysGluLysGlyLysArgGlyGlyGlyAspGluGluProGly 532
Db 1767 AAGCTTCTGAAAGAAACAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1823
Oy 533 GluTyrGlnHisIleHisThrGluSerProGluSerAlaAspGluGluLysGlyProLeuGlu 552
Db 1824 GAGCACTCAGAGTCCAGCCGATTCCTCCGACACCCAGAGAGAGAGAGAGAGAGAGAGAG 1883
Oy 553 SerAlaSerSerProGluGluProGluGluThrThrCysLeuGluLysGlyProLeuGlu 572
Db 1884 TCTGCCTCATCCCTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1943
Oy 573 AlaProGluAspGlyGluAlaGluGluGluThrThrSerAspGlyGluLysArgGlu 592
Db 1944 GTGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2003
Oy 593 GlyIleThrProThrAlaSerPheLysLysMetValThrProLysArgValArgArg 612
Db 2004 GGTGCATCTCCTGGGATCATCTCAAAAAGATGGTGAAGCCCAAGAACCTGTTGAAGCG 2063
Oy 613 ProSerGluSerAspLysGluGluGluGluLysValLysSerAlaThrLeuSerSer 632
Db 2064 CCTTCGAAAGATGAAGAGATGAGCTGAGCAAGAGTCAAGAGCGCTTACTGCTTCC 2123
Oy 633 ThrAspSerThrValSerGluMetGlnAspGluValLysThrValGlyGluGluLys 652
Db 2124 ACCGAGACACAGCCTCTGAAATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2183
Oy 653 ProGluGluProLysArgArgValAspThrSerValSerTTPGluAlaLeuIleCysVal 672
Db 2184 CCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2243
Oy 673 GlySerSerLysLysArgAlaArgLysAlaSerSerSerAspAspGluGlyProArg 692
Db 2244 GATCATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2303
Oy 693 ThrLeuGlyGlyAspSerHisArgAlaGluGluAlaSerLysAspLysGluAlaGlyThr 712
Db 2304 GCAATGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2363
Oy 713 AspAlaValProAlaSerThrGluGluGluAspGlnAlaGluGlySerSerSerProGlu 732
Db 2364 GACGGAGATCTTGTGCTGTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2423
Oy 733 ProAlaGlySerProSerGluGlyGluGlyValSerThrTTPGluSerPheLysArgLeu 752
Db 2424 CAAGCTGGAAGCTTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2483
Oy 753 ValThrProArgLysLysSerLysSerLysLeuGluGluLysValGluAsp-----769
Db 2484 GTACAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2543
Oy 770 ---SerSerValGluGluLysSerThrGluIleGluProSerArgGluGluSerThrVal 788
Db 2544 GGATGTGTGTAAGCATTTCCATCCAGACACTGAACCGGTAAAGAGAGATCTGGGTC 2603

```

QY 789 SerIleLysLysPheIleProGlyArgArgLysLysArgAlaAspGlyLysGlnGluGln 808
DB 2604 TCRAATCAAGAAGTTTATCTCTGGAAGGAAGAAAGGCCAGATGGGAACAAGAACAA 2663
QY 809 AlaThrValGluAspSerGlyProValGluIleAsnGluAspProAsnValProAla 828
DB 2664 GCCCTGTTGAAGACGAGGGCCCAACAGGGCCCAAGATGACTCTGTATGTCGCCGCC 2723
QY 829 ValValProLeuSerGluTyrAsnAlaValGluArgGluLysMetGlu-----Ala 845
DB 2724 GTGGTCCCTCTGTCTGAGTATGCTGTAGAAAGGAGAAATGGAGGCACAGCAAGCC 2783
QY 846 GluGlyAsnThrGluLeuProGlnLeuGlyAlaValTyrValSerGluGluLeuSer 865
DB 2784 CAAAAGCGCAGAGCGCCGAGCAGAGCGCCACTGAGGTGTCCAAAGGAGCTCAGC 2843
QY 866 LysThrLeuValHisThrValSerValAlaValIleAspGlyThrArgAlaValThrSer 885
DB 2844 GAGAGTCAGGTTTCATATGATGGCAGCAGCTGTCTGCTACGGGACGAGCGCTACCAT 2903
QY 886 ValGluGluArgSerProSerTrpIleSerAlaSerValThrGluProLeuGluHisThr 905
DB 2904 ATTGAAGAAAAGGCTCTCTCTTGGATATCTGCTTCACTGACAGAACCTCTTGAACAAGTA 2963
QY 906 AlaGlyGluAlaMetProProValGluGluValThrGluLysAspIleIleAla---Glu 924
DB 2964 GAAGCTGAAGCCGACACTGTTAACTGAGGAGGTATTGAAAGAGAGAAATGTCAGAGAA 3023
QY 925 GluThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHisAspAspMetVal 944
DB 3024 GAACCCCGCCCGTTACTGAACCTCTCCAGAGAACAGAGCGCCCGGCGCACACGGTC 3083
QY 945 ThrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeu 964
DB 3084 GTTAGTGAGCGGAATTGACCCCGAAGCTGTGACAGCTGCAGAACTGCGAGGCCATTG 3143
QY 965 ArgThrGluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSerAla 984
DB 3144 GGTTCGAAGAAGAAACCGAAGCATCTGCTGCTGAAGAGACCACAGAAATGGTGTGAGCA 3203
QY 985 ValSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGluVal 1004
DB 3204 GTCTCCCGATTAAACCGACTCCCGAGACACACAGAGAGGCCACTCCGGTGCAGGAGGTG 3263
QY 1005 GluSerGlyValLeuAspThrGluGluGluArgGlnThrGlnAlaIleLeuGlnAla 1024
DB 3264 GAAGTGGCGTACTGTGACATAGAAGACAGACAGAGCGCGGACTCAAGAGGTCTCTCAGGCA 3323
QY 1025 ValAlaAspLysValLysGluGluSerGlnValProAlaThr----- 1038
DB 3324 GTGGCAGAAAAGTGAAGAGGAATCCAGCTGCTGGCACCGTGGGCCAGAGATGTG 3383
QY 1039 ---GlnThrValGlnArgThrGlySerLysAlaLeuGluLysValGluGluValGluGlu 1057
DB 3384 CTTGAGCTGTGAGAGCA-----GAGCGCAGAAAGACCAGAGAG 3425
QY 1058 AspSerGluValLeuAlaSerGluLysGluLysAspValMetProLysGlyProValGln 1077
DB 3426 CAGGCTGAAGCGTCGGTCTGAAGAAGACACCGATGTAGTGTGAAGTAGATGTCTCAG 3485
QY 1078 GluAlaGlyAlaGluHisLeuAlaGlnGlySerGluThrGlyGlnAlaThrProGluSer 1097
DB 3486 GAGGCAAAAACCTGAGCCTTTTACACAGGGAAGTGTGTGGGCGAGCACCCACCCAGAAAGC 3545
QY 1098 LeuGlu---ValProGluValThrAlaAspValAspHis-----ValAlaThrCys 1113
DB 3546 TTTGAAAAGAGCTCCTCAAGTCACAGAGACATGAGTCCAGTCCAGTGTGTAACCACTTGT 3605
QY 1114 GlnVal-----IleLysLeuGlnGlnLeu---MetGluGlnAlaValAla 1127
DB 3606 CAAGCCGAACCTTAGTCTGGGTAAATAATCAGAGAGATGTGTGAACAGCGCTATCCCC 3665

QY 1128 ProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrProLeuAlaAsp 1147
DB 3666 CTTGACTCTGGTGAACCCCTACAGACAGTGAAGTGTGAAGCACCCTCCCTAGCCGAC 3725
QY 1148 SerAspThrAlaAspGlyThrGlnGlnAspGluThrIleAspSerGlnAspSerLysAla 1167
DB 3726 TTTGACGCACCCAGCACCAAGAGCAGATGTTGTGAATTCATCAGGAGATGAG 3785
QY 1168 ThrAlaAlaValArgGlnSerGlnValThrGluGluGluAlaAlaThrAlaGlnLysGlu 1187
DB 3786 GTGCATCTGTGTA---CACTGAGGGGCACAGAAAGCAGAGCAGTCTCTGCACAGAAAGAG 3842
QY 1188 GluProSerThrLeuProAsnAsnValProAlaGlnGluGluHisGlyGluGluProGly 1207
DB 3843 AGCCCTCCAGCA---CCTTCCAGTTTGTGTTCCAGAAAGAACTAAAGAACCAACAG 3899
QY 1208 ---ArgAspValLeuGluProThrGlnGlnGluLeuThrAlaAlaValProValLeu 1226
DB 3900 ATGGAAGACACTCTAGAGCATACAGATAAAGAGGTCTCAGTGGAACCTGTATCCATTCTG 3959
QY 1227 AlaLysThrGluValGlyGlnGluGlyGluValAspTrpLeuAspGlyGluLysValLys 1246
DB 3960 TCAAGACTGAGGGGACTCAA---GAGGCTGACAGTATGCTGTGAGAAACCAA 4013
QY 1247 -----GluGlnGlnGluValPheValHisSerGly-----ProAsnSer 1259
DB 4014 GACGTACCATTTTTTCAAGGACTTGAGGGGTCTATAGACACAGGCATACAGTCAGTCGG 4073
QY 1260 GlnLysAlaAlaAspValThrTyrAspSerGluValMetGlyValAlaGlyCysGlnGlu 1279
DB 4074 GAAAGGTCACTGAAGTTCCTTAAAGGTGAAGGACAGAGCAAGAGCTGAATGTAAAGAG 4133
QY 1280 LysGluSerThrGluValGlnSer-----LeuSerLeuGluGluGlyGlu 1294
DB 4134 GATGATGCTCTTGAAGTGCAGAGTACGCTAAGTCTCTCCATCCCTCCGCTGGAGAGAG 4193
QY 1295 MetGluThrAspValGluLysGluLysArgGluThrLysProGluGlnValSerGluGlu 1314
DB 4194 ATGGTAGTTCAAGTCAAGAGGAGAAACAGAAAGCAGAGCAACCCATGTGAATGAAGAG 4253
QY 1315 Gly---GluGlnGluThrAlaAlaProGluHisGluGlyThrTyrGlyLysProValLeu 1333
DB 4254 AGCTTGACGACGAAACAGCTGTACCGTATCTGAAGAGGTCAAGTAAAGCAGCTCCTCCAG 4313
QY 1334 ThrLeuAspMetProSerSerGluArgGlyLysAlaLeuGlySerLeuGlyGlySer--- 1352
DB 4314 ACAGTGAATGTGCCCATCATAGATGGGCAAGAGAGTCAAGAGTTTGAAGAAGCCCT 4373
QY 1353 ProSerLeuProAspGlnAspLysAlaGlyCysIleGluValGlnValGlnSerLeuAsp 1372
DB 4374 CTTCCCTGCTAGGTCAAGAGGAGGAGTATGCACCAAAATTCAGTTTCAGAGCTCTGAG 4433
QY 1373 ThrThrValThrGlnThrAlaGluAlaValGluLysValIle-----GluThrValVal 1390
DB 4434 GCATCAATTCATCTTAACGCGCTGCAGAGGAGGAAAGGTCTTAGGAGAAATTCGCAAC 4493
QY 1391 IleSerGluThrGlyGluSerProGluCysValGlyAlaHisLeuLeuProAlaGluLys 1410
DB 4494 ATTTAGAAACAGGTGAACAGTTGGAGCTGCAGGTGCACATTTAGTTCTGGAAAGAA 4553
QY 1411 SerSerAlaThrGlyGlyHisTrpThrLeuGlnHisAlaGluAspThrValProLeuGly 1430
DB 4554 TCTCTGAAAAAATGAAGACTTTGCGCTCATCCAGGGGAGAGTCTGTGCCACAGGG 4613
QY 1431 ProGluSerGlnAlaGluSerIleProIleIleValThrProAlaProGluSerThrLeu 1450
DB 4614 CCGCACTGTGCGCAAAATCGACCCAGTGTATCTCTACTACTACCAAGAAAGGCTTA 4673
QY 1451 HisProAspLeuGlnGlyGluIleSerAlaSerGlnArgGluArgSerGluGluGluAsp 1470
DB 4674 AGTTCCGACTCTGGAAGGAGAAACCAACATCATCTGAAGTGAAGTCAGATGATCGAT 4733
QY 1471 LysProAspAlaGlyProAspAlaAspGlyLysGluSerThrAlaIleGluLysValLeu 1490


```

Db 4734 GAGCAGGTTGCTTCCAGGAGTC-----AAAGTAGCTGTGCAATTGAGGATTTA--- 4784
Oy 1491 LysAlaGluProGlu-----IleuGluLeuGluSerLysSerAsnLysIleValLeu 1508
Db 4785 -----GAGCCGAAATGGGATTTTGAACTTGAGAACCAAAAGCAGTAACTGTGCCAA 4838
Oy 1509 AasnValIleGluThrAlaValAspGluPheAlaArgThr---GluThrAlaProGluThr 1527
Db 4839 AACATCATCTCACAACAGCCCTTGACCGTTGTGACGTACAGAAACACACCCGAAATG 4898
Oy 1528 HisAlaIAspSerSerGluThrGluValProAlaCysArgLeuAspSerArgGluProAsn 1547
Db 4899 TTGACGCTGAGTTACAGACACAGCTCAGTATAAAGCTGACAGCCAGACGCTGGA 4958
Oy 1548 ArgCysTrpThrLys-----MetLysAspAla-LysMetLysHisProVa 1562
Db 4959 CAGGAAACGAGAAAGAAAGAGAGAACTCAGGCCTCTGCACAGATGAAACCAATT 5018
Oy 1562 LProGluProArgGlu-----AspLeuGluValLeuThrValLe 1575
Db 5019 ACTTCAGCCAAAGAGAGAGTCAAGTCAACCGCAGTGGGACAAAGCATTCTGTATTTC 5078
Oy 1575 uGluAlaTrpAlaGluProArgLysCysLeuProArgLeuGluLeuAlaProVal 1594
Db 5079 AAAGACATGATGAGACCTCAGAAAGAACCATGACTGTGAGGTAGAGGTTCCACTG 5136

RESULT 7
AAA74903
ID AAA74903 standard; cDNA; 6605 BP.
XX
AC AAA74903;
XX
DT 26-JAN-2001 (first entry)
XX
XX Human gravin coding sequence.
DE
XX Human; gravin; PKA RII binding site; myasthenia gravis;
KW kinase anchoring protein; CAMP dependent protein kinase; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 192..5534
FT FT /*tag= a
FT FT /product= "gravin"
XX
XX US6090929-A.
XX
XX 18-JUL-2000.
XX
XX 19-DEC-1997; 97US-0994570.
XX
XX 19-DEC-1996; 96US-0769309.
XX
XX (UYOR-) UNIV OREGON HEALTH SCI.
XX
XX KLauck TM, Scott JD, Nauert JB;
XX
XX WPI; 2000-523763/47.
XX
XX P-PSDB; AAB15380.
XX
XX Novel polynucleotides useful for detecting gravin in patients suffering
XX from Myasthenia gravis encodes CAMP-dependent protein kinase-binding
XX polypeptide and protein kinase C-binding polypeptide of gravin -
XX
XX Example 1; Column 23-36; 34bp; English.
XX
XX The present sequence is the coding sequence for human gravin. GravIn is
XX an A-kinase anchoring protein (AKAP) which is involved in the
XX localisation of cAMP dependent protein kinase A (PKA) via interactions
XX between the RII binding region and the PKA regulatory subunit RII. GravIn
XX is also an antigen found in myasthenia gravis sufferers, and it is

```

```

CC thought that antibodies to it may be useful in modulating the binding of
CC PKA, and thus aid in the treatment of the disease. This sequence was
CC isolated by first screening a human umbilical vein endothelial cell cDNA
CC library with serum from a myasthenia gravis patient, and then searching a
CC human heart cDNA library for sequences resembling the isolated sequence.
CC This was done because the first sequence obtained was shown to be shorter
CC than the full length coding sequence.
XX
SQ Sequence 6605 BP; 2091 A; 1416 C; 1826 G; 1272 T; 0 other;

Alignment Scores:
Pred. No.: 1,72e-211 Length: 6605
Score: 4281.00 Matches: 965
Percent Similarity: 68.69% Conservative: 189
Best Local Similarity: 57.44% Mismatches: 409
Query Match: 53.03% Indels: 117
DB: 21 Gaps: 37

US-09-902-432-4 (1-1596) x AAA74903 (1-6605)
Oy 1 MetGlyAlaGlySerSerThrGluGlnArgSerProGluGln---ProAlaGlySerAsp 19
Db 192 ATGGGCGCGGGAGGCTCCACCGAGCAGCGCAGCCGCGAGCAGCCGCGAGGGAGCTCC 251
Oy 20 ThrProSerGluLeuValLeuSerGlyHisGlyProAlaAlaGluAlaSer---GlyAla 38
Db 252 ACGCGGCTGAGCCGAGCCGAGCGCGCGGCCCTCTGCGCGAGGCGCCGAGACACC 311
Oy 39 AlaGlyAspProAlaAspAla-----AspProAlaThrLysLeuProGluLysAsnGly 56
Db 312 ACCGGGGAACCCCGCATGCTGCTCCGAGACCCCGGCAACAGCTCTCTACAGAAGATGT 371
Oy 57 GlnLeuSerSerValAsnGlyValAlaGluGlnGlyAspValHisValGlnGluGln 76
Db 372 CAGCTGTCCACCATCAATGCGGTAGCTGAGCAAGATGAGCTCAGCTCCAGAGGGTGAC 431
Oy 77 GlnGluGlyGln-----GlnGlu 82
Db 432 CTAAATGGCCGAAAGAGCCCTGAAACGCTCAAGAGCCCTTAACAGCCAGAGGAAAGA 491
Oy 83 GluValAlaAspGluAspValGlyGlnArgGluSerGluAspValArgLysAspArg 102
Db 492 GAAAGCATTTGTACCGAGAGTTGGACAGAGAGACTTGTAAGATGTAGGAAAGACATCC 551
Oy 103 ValGluGluMetAlaIAsnSerThrAlaValGluAspLleThrLysAspGlyGlnGlu 122
Db 552 GATAAAGAGATGGCTAAGTCAAGTCAAGCGGTGTTCACGACATCACAGATGATGGCAGAG 611
Oy 123 GluThrSerGluIleIleGluGlnIleProAlaSerGluAsnAsnValGluGluMetVal 142
Db 612 GAGAACCGAAT--ATCGAACAGATTCCTTCAGAAAGCAATTTGAAAGAGCTAAACA 668
Oy 143 GlnProAlaGluSerGlnAlaAsnAspValGlyPheLysLysValPheLysPheValGly 162
Db 669 CAACCCACTGAGTCCCAAGGCTATATGATTTAAGAGGTITTAAGTTTGTGGC 728
Oy 163 PheLysPheThrValLysLysAspLysAsnGluLysSerAspThrValGlnLeuLeuThr 182
Db 729 TTTAATTCACTGTGAAAAAGGATTAAGACAGAGACCTGACACTGTCCAGCTACTCACT 788
Oy 183 ValLysLysAspGluGlyGluGlyLysLysLysValAspValGlyAlaGlyAspHisGlnGlu 202
Db 789 GTGAAGAAAGATGAAGGAGGAGGAGCA-----GCAGGGGCTGGCGACCAACAGAGAC 839
Oy 203 ProSerValGluThrAlaValGluSerAlaSerLysGluSerGluLeuLysGlnSer 222
Db 840 CCCAGCTT-----GGGCTGGAAGACAGCATCCAAAGAAAGCAACCCAAACATCT 893
Oy 223 ThrGluLysGlnGluGlyThrLeuLysGlnGlnGlnSerSerThrGluIleProLeuGln 242
Db 894 ACAGAGAAACCGGAGAGAGACCTGAAAGCGTGAAGCAAGCAGCAAGAAATTTCTCCCCCA 953
Oy 243 AlaGluSerAspGlnAlaAlaGluGluGluAlaLysAspGluGlyGluGluLysGlnGlu 262

```


Db 954 GCCGAATCTGCCAAGCAGTG- --GAGAAATGCAAGGAGGAGGAGGAGAAACAAGAA 1010
Qy LysGluProThrLysSerProGluSerProSerProValAenSerGluThrThrSer 282
Db 1011 AAGAACCTAGCAAGTCTGCAGAAATCTCCGACTAGTCCCGTGACCAAGGAGATCA 1070
Qy 283 SerPheLysLysPhePheThrHisGlyTTPalaGlyTTPalaGlyLysThrSerPheLys 302
Db 1071 ACCTTCAAAAATTTCTTCACTCAAGGTTGGCGCGCTGGCGCAAAAAGACCAAGTTTCAGG 1130
Qy 303 LysSerLysGluAspLeuThrAlaGluLysArgLysGluGlnGluAlaGluLys 322
Db 1131 AAGCCGAAGGAGGATGAAGTGAAGCTTTCAGAGAAGAAAGGAACAGGACCCAGAAAAA 1190
Qy 323 ValAspGluGluGluLysGluLysThrGluProAlaSerGluGlu- ----- 337
Db 1191 GTAGACACAGAGAAGACGGAAGGACAGAGTTTGCCTCCGAGAAACTGACCGCTCCGAG 1250
Qy 338 -----GlnGluProAlaGluAspThrAspGlnAlaArgLeuSerAlaAspTyr 353
Db 1251 CAAGCCACCCACAGGAGCGCGCAAAAGTGCCACAGCGCCCGGTTATCAGCTGAAATAT 1310
Qy 354 GluLysValGluLeuProLeuGluAspGlnValGlyAspLeuGluAlaSerSerGluGlu 373
Db 1311 GAGAAAGTTGAGCTGCCCCCTCAGAGGAGCAAGTCAAGTGGCTCGCAGGACCTTCTGAAGAG 1370
Qy 374 LysCysAlaProLeuAlaThrGluValPheAspGluLysMetGluAlaHisGln- --Glu 392
Db 1371 AAACCTGCTCGTTGGCGACAGAGTGTTCATGAGAAATAGNAGTCCACCAAGAGAG 1430
Qy 393 ValValAlaGluValHisValSerThrValGluLysThrGluGluGlnGlyGly 412
Db 1431 GTTGTGGCGAAGTCCACGTCAGCACCGCTGAGGAGGAGAACCGAAGAGCAG- ----- 1481
Qy 413 GlyGluAlaGluGlyGlyValValValGluGlyThrGlyGluSerLeuProGluLys 432
Db 1482 -----AAAAAGGAGTGAAGAAACACAGCAGGGTCTGTGCCAGCTGAAGAA 1526
Qy 433 LeuAlaGluProGlnGluValProGlnGluAlaGluProAlaGluGluLeuMetLysSer 452
Db 1527 TTGTTGGAATGGATGAGACCTCAGAGACCGCAACTGCCAAGGAGCTGGTGAAGCTC 1586
Qy 453 ArgGluMetCysValSerGlyGlyAspHisThrGlnLeuThrAspLeuSerProGluGlu 472
Db 1587 AAAAAACGTGTGTTCCGAGAGGACCTACACAGGAGCTGACCTCAGTCTCTGTATGAG 1646
Qy 473 LysThrLeuProLysHisProGluGlyLysValSerGluValGluMetLeuSerSerGln 492
Db 1647 AAGTGTCTGTCCAAACCCCGAAGGCGTGTGTGAGTGAGGTGGAATGCTGTATCATCAG 1706
Qy 493 GluArgLysValGlnGlySerProLeuLysLysLeuPheSerSerSerGlyLeuLys 512
Db 1707 GAGAGATGAAGTGCAGGAGAGTCCACTAAGNAGCTTTTACCAGCACTGGCTTAAAA 1766
Qy 513 LysLeuSerGlyLysLysGlnLysGlyLysArgGlyGlyGlyAspGluGluProGly 532
Db 1767 AAGCTTTCTGAAAGAAACAAGAAAGGAAAAA- --GGAGGAGGAGACGAGGAATCAGGG 1823
Qy 533 GluTyrGlnHisLysHisThrGluSerProGluSerAlaAspGluGlnLysGlyGluSer 552
Db 1824 GAGCACACTCAGGTTCCAGCGCATTTCTCCGACACGAGGAGGAGCAAAAAGGCGAGAGC 1883
Qy 553 SerAlaSerSerProGluGluProGluThrThrCysLeuGluLysGlyProLeuGlu 572
Db 1884 TCTGCTCATCCCTGAGGAGCCCGAGAGATACAGTGTCTGAAAAAGGGCTTAGCCGAG 1943
Qy 573 AlaProGlnAspGlyGluAlaGluGluGlyThrThrSerAspGlyGluLysLysArgGlu 592
Db 1944 GTGCAGCAGGATGGGAAGCTGAAGAGGAGCTTCTCCGATGGAGAGAAAAAAGAGAA 2003
Qy 593 GlyLysThrProThrAlaSerPheLysMetValThrProLysLysArgValArg 612

Db 2004 GGTGTCACTCCCTCGGCGATCAATTCAAAAAGATGGTGCAGCCCAAGAGCGGTGTAGACGG 2063
Qy 613 ProSerGluSerAspLysGluGluGluLeuGluLysValLysSerAlaThrLeuSerSer 632
Db 2064 CCTTCGGAAGTGATAAAGAAGATGAGTGGACAAGGTCAAGAGCGCTACTCTGTCTTCC 2123
Qy 633 ThrAspSerThrValSerGluMetGlnAspGluValLysThrValGlyGluGluGlnLys 652
Db 2124 ACCGAGAGCACAGCCTCTGAAATGCAAGAAGAAATGAAAGGGAGCGTGGGAAGACCAAG 2183
Qy 653 ProGluGluProLysArgValAspThrSerValSerTrpGluAlaLeuLysVal 672
Db 2184 CCGAAGAACCAAGCCAGGTGGATACCTCAGTATCTTTGGGAAGCTTTAATTTGTGTG 2243
Qy 673 GlySerSerLysLysArgAlaArgLysAlaSerSerSerAspGluGlyGlyProArg 692
Db 2244 GGATCATCAAGAAAAAGAGCAAGGAGAGGTCTCTTCTGATGAGGAAGGGGACCAAAA 2303
Qy 693 ThrLeuGlyGlyAspSerHisArgAlaGluGluAlaSerLysAspLysGluAlaGlyThr 712
Db 2304 GCAATGGGAGGAGACCCACAGAAAGCTGATGAGCGCGGAAACACAAAGAGACGGGACA 2363
Qy 713 AspAlaValProAlaSerThrGlnGluGlnAspGlnAlaGlnGlySerSerProGlu 732
Db 2364 GACGGATCTTCTGCTTCCCAAGAACATGATCCAGGGCAGGGAAGTTCTCTCCCGGAG 2423
Qy 733 ProAlaGlySerProSerGluGlyGlyValSerThrTrpGluSerPheLysArgLeu 752
Db 2424 CAAGCTGGAAGCCCTACCGNAGGGAGGGCTTTCACCTGGGAGTCATTTAAAGGTTA 2483
Qy 753 ValThrProArgLysLysSerLysLeuGluGluLysAlaGluAsp- ----- 769
Db 2484 GTCACGCCAAGAAAAAATCAAGTCCAAGTGGAAAGAGAAAAAGCGAAGACTCATAGCT 2543
Qy 770 ---SerSerValGlnLeuSerThrGluLysGluProSerArgGluGluSerTrpVal 798
Db 2544 GGGTCTGGGTGAGAACATTCCTCCTCAGACACTGAAACCCGGTAAAGAAAGAAATCCTGGGTC 2603
Qy 789 SerLysLysPheLysProGlyArgArgLysLysArgAlaAspGlyLysGlnGluGln 808
Db 2604 TCATCAAGAGTTTATTCCTGGACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2663
Qy 809 AlaThrValGluAspSerGlyProValGluLysLeuAsnGluAspProAsnValProAla 828
Db 2664 GCCCTGTTGAAGACGACGAGGCGCAACAGGGGCCAACGAGATCACTCTGTATGTCGCGCC 2723
Qy 829 ValValProLeuSerGluThrAsnAlaValGluArgGluLysMetGlu- -----Ala 845
Db 2724 GTGGTCCCTCTGTCTGATGATGCTGTAGAAAGGGAGAAAAATGGAGGCACAGCAAGCC 2783
Qy 846 GlnGlyAsnThrGluLeuProGlnLeuLeuGlyAlaValTyrValSerGluGluLeuSer 865
Db 2784 CAAAAAGGCGCAGAGCAGCGCCGAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2843
Qy 866 LysThrLeuValHisThrValSerValAlaValLysAspGlyThrArgAlaValThrSer 885
Db 2844 GAGAGTCAGGTTTATGATGAGCAGCAGTGTCTGTCAGCGGACGAGGAGGAGGAGGAGGAG 2903
Qy 886 ValGluGluArgSerProSerTrpLysSerAlaSerValThrGluProLeuGluHisThr 905
Db 2904 ATTGAAGAAGGTTCTCTTCTGGATATCTGCTTCAAGGAGGAGGAGGAGGAGGAGGAGGAG 2963
Qy 906 AlaGlyGluAlaMetProProValGluGluValThrGluLysAspLysLeuAla- --Glu 924
Db 2964 GAAGCTGAAGCCCACTGTTTAACTGAGGAGGATTTGGAAGAGAGAGGAGGAGGAGGAGGAG 3023
Qy 925 GluThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHisAspAspMetVal 944
Db 3024 GAACCCCCACCGGTTTACTGAACCTCTGCCAGAGAACAGAGAGGAGGAGGAGGAGGAGGAG 3083
Qy 945 ThrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeu 964
Db 3084 GTTAGTGAGGCGGAATTTGACCCCGGAAAGCTGTGACAGCTGCAGAAACTGCAGGGCCATTG 3143

```

QY 965 ArgThrGluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSerAla 984
Db 3144 GGTTCGGAAGAAAGAAACCAACATCTGCTGTGAAGAAACCAACAAATGGTGTACAGA 3203
QY 985 ValSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGluVal 1004
Db 3204 GTCTCCAGTTAACGACTCTCCAGACACACAGAGAGGCCACTCCGGTGCAGAGGTG 3263
QY 1005 GluSerGlyValLeuAspThrGluGluGluValArgGlnThrGlnAlaIleLeuGlnAla 1024
Db 3264 GAAGGTGGGTACCTGACATAGAAAGCAAGAGAGCGGACTCAAGAGTCTCCAGCA 3323
QY 1025 ValAlaAspValValGluGluGluSerGlnValProAlaThr----- 1038
Db 3324 GTGGCAGAAAAAGTGAAGAAAGAAATCCAGCTGCGCCAGCCGGTGGCCAGAAAGTGTG 3383
QY 1039 ---GlnThrValGlnArgThrGlySerLysAlaLeuGluValGluValGluGlu 1057
Db 3384 CTTACGCTGTGTCAGAGACA-----GAGCCAGAAAGACCAAGAGAG 3425
QY 1058 AspSerGluValLeuAlaSerGluLysGluLysAspValMetProLysGlyProValGln 1077
Db 3426 CAGGCTGAAGCGTCCGGGTCTGAAGAAAGAGACGAGATGTAGTGAAGTAGATGCTCAG 3485
QY 1078 GluAlaGlyAlaGluHisLeuAlaGlnGlySerGluThrGlyGlnAlaThrProGluSer 1097
Db 3486 GAGGCAAAAACTGAGCGCTTTTACACAAAGGAAAGTGTGGGCGACACCCCAAGAAC 3545
QY 1098 LeuGlu---ValProGluValThrAlaAspValAspHis-----ValAlaThrCys 1113
Db 3546 TTGGAAGAAAGCTCTCACTCACTACAGAGACATAGACTCCAGTGAAGTTAACCACTTGT 3605
QY 1114 GlnVal-----IleLysLeuGlnGlnLeu---MetGluGlnAlaValAla 1127
Db 3606 CAGCGCAAAACCTTAGCTGGGGTAAATATCACAGAGATGTGATGGAACAGGCTATCCCC 3665
QY 1128 ProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrProLeuAlaAsp 1147
Db 3666 CCGTACTCGGTGAAACCCCTTACAGACATGTAGACTGATGAAAGCCCCCGTAGCCGAC 3725
QY 1148 SerAspThrAlaAspGlyThrGlnGlnAspGluThrIleAspSerGlnAspSerLysAla 1167
Db 3726 TTGAGGCCACCGACCAACCCAGAAAGACGAGATTGTGAAATCCATGAGAGAAATGAG 3785
QY 1168 ThrAlaAlaValArgGlnSerGlnValThrGluGluGluAlaAlaThrAlaGlnLysGlu 1187
Db 3786 GTGCATCTGTGA---CCAGTCAAGGGGCACAGAGAGGCAAGTCTCTCACAGAAAGAG 3842
QY 1188 GluProSerThrLeuProAsnAsnValProAlaGlnGluGluHisGlyGluGluProGly 1207
Db 3843 AGCGCTCCAGCA---CCTTCCAGTTTGTGTTCCAGAAAGAACTAAACAAACATCAAG 3899
QY 1208 ---ArgAspValLeuGluProThrGlnGlnGluThrAlaAlaAlaValProValLeu 1226
Db 3900 ATGGAAGACACTTACAGCATACATAGAGAGGTCTCAGTGAACCTGATTCATTCTG 3959
QY 1227 AlaLysThrGluValGlyGlnGluGluValAspTrpLeuAspGlyGluLysValLys 1246
Db 3960 TCAGAGACGACGAGGGGACCTCA-----GAGCGTCAACAGATGCTGATGAGAAACAA 4013
QY 1247 -----GluGluGlnGluValPheValHisSerGly-----ProAsnSer 1259
Db 4014 GACGTACCAATTTTTCAGAGACTTGAAGGGGTCTATGACACAGGCGCTAAACAGTCACTCG 4073
QY 1260 GlnLysAlaAlaAspValThrTrpAspSerGluValMetGlyValAlaGlyCysGlnGlu 1279
Db 4074 GAAAGAGTCACTGAAGTGTCCCTTAAAGTGAAAGGACAGAGAAACCTGAATGTAAGAAAG 4133
QY 1280 LysGluSerThrGluValGlnSer-----LeuSerLeuGluGluGluGlu 1294
Db 4134 GATGATGCTCTTGAATGACAGACTCACGCTAAGTCTCTCCATCCCCCGTGAAGAGAGAG 4193

```

```

QY 1295 MetGluThrAspValGluLysGluLysArgGluThrLysProGluGlnValSerGluGlu 1314
Db 4194 ATGTGATGTTCAAGTGCAGAAAGGAGAGAAACAGACAGAGGCCACCATGTATGATGAAGAG 4253
QY 1315 Gly---GluGlnGluThrAlaAlaProGluHisGlyGluThrTrpGlyLysProValLeu 1333
Db 4254 AAGCTTAGACAGAAACAGCTGTTAACGCTATCTGAAGAGGTCAATAGACACTCTCCAG 4313
QY 1334 ThrLeuAspMetProSerSerGluArgGlyLysAlaLeuGlySerLeuGlyGlySer-- 1352
Db 4314 ACAGTGAATGTGCCCATCATGATGAGGCGCAAGAAAGACACACAGCTTTGGAAGAGCCCT 4373
QY 1353 ProSerLeuProAspGlnAspLysAlaGlyCysIleGluValGlnValGlnSerLeuAsp 1372
Db 4374 CTTCCCTGCTAGGTCAGAGAGAGGAGGTATGACACAAATTCATCAAGTCAAGCTGAG 4433
QY 1373 ThrThrValThrGlnThrAlaGluAlaValAlaGluLysValIle-----GluThrValVal 1390
Db 4434 GCATCATTCACCTTACACAGCGGCTGCAGAGAGAGAAAGCTTAAAGAGAACTGCCAAC 4493
QY 1391 IleSerGluThrGlyGluSerProGluCysValGlyAlaHisLeuLeuProAlaGluLys 1410
Db 4494 ATTTTGAAGAACAGGTGAACCGTTGAGCTTCAGGTCGACATTTAGTTTCTGGAAGAGAA 4553
QY 1411 SerSerAlaThrGlyHisIleTrpThrLeuGlnHisAlaGluAspThrValProLeuGly 1430
Db 4554 TCCTTGAAAAAATGAAGACTTTCGCCCTCATCCAGGGGAAAGTCTGCCACAGGG 4613
QY 1431 ProGluSerGlnAlaGluSerIleProIleValThrProAlaProGluSerThrLeu 1450
Db 4614 CCCGACTGTACAGCAAAATTCACACACAGATAGTACTGTCTACTACCAAGAAAGGCTTA 4673
QY 1451 HisProAspLeuGlnGlyGluIleSerAlaSerGlnArgGluArgSerGluGluGluAsp 1470
Db 4674 AGTTCCGACCTGGAAGGAGAAACCATCATCAGTGAAGTGAAGTCAAGTGAAGTGCAT 4733
QY 1471 LysProAspAlaGlyProAspAlaAspGlyLysGluSerThrAlaIleGluLysValLeu 1490
Db 4734 GAGCAGGTGCTTGCACAGAGGTC-----AAAGTGAATGTAGCAATTGAGAGATTTA-- 4784
QY 1491 LysAlaGluProGlu-----IleLeuGluLeuGluSerLysSerAsnLysIleValLeu 1508
Db 4785 -----GAGCTGGAAGAAATGGGATTTTGGAACTTGAGACCAAAAGAGATTACTGTCCA 4838
QY 1509 AsnValIleGlnThrAlaValAspGlnPheAlaArgThr---GluThrAlaProGluThr 1527
Db 4839 AACATCATCCAGACAGCCGTTGACACAGTTTGTACGTACAGAAAGAAACAGCCAGAAATG 4898
QY 1528 HisAlaTrpAspSerGlnThrGlnValProAlaCysArgLeuAspSerArgGluProAsn 1547
Db 4899 TTGACGTGTGAGTTACACACACAGCTCACTGATTAAGCTGACAGCCAGAGACGCTGGA 4958
QY 1548 ArgCysTrpThrLys-----MetLysAspAla-LysMetLysHisProVal 1562
Db 4959 CAGGAAGAGGAAAGAGAGAGAGAGAACCTCAGCGCTTGACAGAGATGAAGAACCAATT 5018
QY 1562 IProGlnProArgGlu-----AspLeuGlnValLeuThrValLe 1575
Db 5019 ACTTCAGCCAAAGAGAGAGTCAAGTCAACCCAGTGGGCAAGCAATTCGATGATTTCC 5078
QY 1575 uGluAlaTrpAlaGlnProArgLysCysLeuProArgLeuGlnLeuLysAlaProVal 1594
Db 5079 AAAGCATGAGTGAAGCTCAGAAAGAACATGACTGTTGAGTGAAGTTCACCTG 5136

```

RESULT 8
 AAS85205
 ID AAS85205 standard; cDNA; 6886 BP.
 AC AAS85205;
 AC AAS85205;
 DT 13-FEB-2002 (first entry)
 XX DNA encoding novel human diagnostic protein #21009.

OY	392	uValValaIaGIuValaIHiVaIseThrValGIuIysthGIuGIuGIuGIuGIuGIuGI	412
Db	1434	GGTTGTGGCCGAAGTCCACGCTCAGACCCGTGGAGGAGGAAACCGAAGGACAG-----	1485
OY	412	YGIYGIuIaGIuGIuGIuValValaGIuGIuGIuGIuGIuGIuGIuLeuProProGIuIy	432
Db	1486	-----AAACGAGGGGGAGAAACAGCAGGGGTCTGTGCCAGCTGAAGA	1529
OY	432	sLeuAlaGIuProGIuGIuValProGIuGIuIaGIuProAlaGIuGIuLeuMetLysSe	452
Db	1530	ATTGGTTGGAAATGGAATGCACAACTCAGAGAGCCGAACCTGGCCAGAGAGCTGTGAAGCT	1589
OY	452	rArgGIuMetCyVaIseGIuGIuYAspHisThrGIuLeuThrAspLeuSeProGIuGI	472
Db	1590	CAAAAGAAACGTGTGTTCCGAGAGGAGCCCTACACAGGAGCTGACCTCAGCTCGATGA	1649
OY	472	uLYeThrLeuProLYsHisProGIuGIuLYIleValSerGIuValaGIuMetLeuSeSerGI	492
Db	1650	GAAAGTGTGTCCAAACCCCCCAAGGCGTTGTGAATGAGTGGAAATGCTGTATCA	1709
OY	492	nGIaArgLIeLYsValaGIuGIuSeProLeuLYsLYsLeuPheSeSeSerGIuLYeLY	512
Db	1710	GGAAGAAATGAAGTGCAGGGAATCCACTTAAGAGCTTTTATCAGACACGTGGTTAA	1769
OY	512	sLYsLeuSeSerGIuLYsLYsGIuLYsGIuLYsArgGIuGIuGIuYAspGIuGIuProGI	532
Db	1770	AAAGCTTTCTGAAAGAAACGAAGGAAAGA---GGAGAGGAGACAGAGAAATCAG	1826
OY	532	YGIuTYrGIuHiAlIeHisThrGIuSeSerProGIuSeSerAlaAspGIuGIuLYsGIuSe	552
Db	1827	GGAGACACCTCAGGTTCCAGCCGATTCCTCGGACACCCAGAGAGGACAAAGGCGAAG	1886
OY	552	rSerAlaSeSerProGIuGIuProGIuGIuThrThrCysLeuGIuLYsGIuProLeuGI	572
Db	1887	CTTGCTCTCATCCCTCAGAGGCCCGAGAGATCACTGTCTGGAAAAGGCTTAAGCCA	1946
OY	572	uAlaProGIuAspLYsGIuIaGIuGIuGIuThrThrSerAspGIuLYuLYsArgGI	592
Db	1947	GGTGACGACGAGATGGGAGAGCTGAAGAGAGCTACTTCGATGGAGAGAAAAAGAGA	2006
OY	592	uGIuLYIleThrProTYrAlaSerPheLYsLYsMetValThrProLYsLYsArgValArg	612
Db	2007	AGGTCTCATCCCTCGGACATCATTCATAAAAGAGTGAAGGCCCAAGAGCGTGTAGACG	2066
OY	612	GProSeSerGIuSeSerLYsGIuGIuGIuLeuGIuLYsValLYsSerAlaThrLeuSeSe	632
Db	2067	GCTTTCGAAAGTATTAAGAAGTGAAGCTGGACAAGGTCGCAAGAGGCTACTTCTTC	2126
OY	632	rThrAspSerThrValSerGIuMetCysAspGIuValLYsThrValaGIuGIuGIuLY	652
Db	2127	CACCGAGAGCACACCTCTGTAAGACAGAGAAAGAAAGGAGAGCGTGAAGAGCCAA	2186
OY	652	sProGIuGIuProLYsArgArgValAspThrSerValSerTPGIuAlaLeuLYsCYsVa	672
Db	2187	GCCGGAAGAAACAAAGCGCAAGGTGATCCTCAGATCTTGGGAAGCTTAAATTGTGT	2246
OY	672	LYsSerSerLYsArgArgValaArgLYsAlaSerSerSerAlaAspAspGIuGIuProAr	692
Db	2247	GGGATCACCCAAAGAAAGAGCAGAGAGGGGTCTCTTCTGATGAGAGAAAGGGGACCAA	2306
OY	692	GThrLeuGIuLYsAspSerHisArgAlaGIuGIuIaSerLYsAspLYsGLuAlaGIuTh	712
Db	2307	AGCAATGGAGAGAACCCACCAAGAAAGCTATGAGGCCCGGAAAAGACAAAGACGGGAC	2366
OY	712	rAspAlaValProAlaSerThrGIuGIuGIuAspGIuAlaGIuGIuLYsSerSerProGI	732
Db	2367	AGACGGGATCTCTGTGTTCCCAAGAACTGATCCAGGGCAGGGAAGTTCTCCCGGA	2426
OY	732	uProAlaLYsSerProSerGIuGIuGIuGIuValaSerThrTrpGIuSerPheLYsArgLe	752
Db	2427	GCAAGCTGGAAAGCCCTTCGAAAGGGAGGGCGTTTCCACTGGAGATCATTTAAAAAGTT	2486

[illegible]

```

Db 3549 GCITTTGAAAAAGCTCTCAAGTACAGAGACATAGAGTCCAGTGCAGCTTGTAAACCACTT 3608
Qy 1113 ysGlnVal-----IleLysLeuGlnGlnLeu---MetGluGlnAlaVala 1127
Db 3609 GTCAAGCGGAACCTTAGCTGGGTAAATACACAGGAGATGGTGCATGCAACAGGCTATCC 3668
Qy 1127 laProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrProLeuAlaA 1147
Db 3669 CCCCTGATCGGTGGAACCCCTACAGACAGTGCAGACTGATGTAAGCAAGCACCCCGTAGCCG 3728
Qy 1147 spSerAspThrAlaAspGlyThrGlnGlnAspGluThrIleAspSerGlnAspSerLysA 1167
Db 3729 ACTTTGACCGACCCAGGCACAAACCCAGAAAGACGAGATGTGGAAATCCATGAGGAGAAATG 3788
Qy 1167 laThrAlaAlaValArgGlnSerGlnValThrGluGluGluAlaAlaThrAlaGlnLysG 1187
Db 3789 AGTTCGATCTGTGATCCAGTCAGGGGSCACAGAGCAGAGGCGATTCCTGCACAGAAAG 3848
Qy 1187 luGluProSerThrLeuProAsnAsnValProAlaGlnGluGluHisGlyGluGluProG 1207
Db 3849 AGAGGCTCCAGCA---CCTTCAGTGTTCCTCCAGGAAGAACTTAAAGAAACAATCAA 3905
Qy 1207 ly---ArgAspValLeuGluProThrGlnGlnGluLeuThrAlaAlaValProValL 1226
Db 3906 AGATGGAAGACACTCTAGAGCATACAGATAAGAGAGGTGTCAGTGGAAACTGTATCCATTC 3965
Qy 1226 euAlaLysThrGluValGlyGlnGluGlyGluValAspTrpLeuAspGlyGluLysValL 1246
Db 3966 TGTCAAGACTGAGGGGACTCAA-----GAGCGTGACCATGTATGTCAGAGAAACCA 4019
Qy 1246 ys-----GluGluGlnGluValPheValHisSerGly-----ProAsnS 1259
Db 4020 AAGACGTACCACTTTTCGAAGGACTTGAGGGGTCTATAGACACAGGCATTAACAGTCAGTC 4079
Qy 1259 erGlnLysAlaAlaAspValThrTyrraspsergluvalmetglyvalalaglycysgln 1279
Db 4080 GGGAAAAGGTCACCTGAAGTTGCCCTTAAAGGTGAAGGACAGAGAAAGCTGAAATGTAAAA 4139
Qy 1279 lulyserGluSerThrGluValGlnSer-----LeuSerLeuGluGluGlyG 1294
Db 4140 AGGATGATGCTCTTGAACTGCAGAGTCACGTAAGTCTCCATCCCTCCGTCGAGAGAG 4199
Qy 1294 luMetGluThrAspValGluGlyLysArgGlu---ThrLysProGluGlnValserG 1313
Db 4200 AGATGGTAGTTCAGTTCGAAAGGAGGAATACCAAGTAGCAGAGCCCAACCATGTGAATG 4259
Qy 1313 luGluGly---GluGlnGluThrAlaAlaProGluHisGlyGlyThrTyrrGlyLysProV 1332
Db 4260 AAGAGAACTTGAGCGACGAAACAGCTGTTACCGTATCTGAAGAGGTCAGTAAGCAGCTCC 4319
Qy 1332 alLeuThrLeuAspMetProSerSerGluArgGlyLysAlaLeuGlySerLeuGlyGlyS 1352
Db 4320 TCCAGACAGTGAATGTGCCATCATAGATGGGCAAGAGAGTCAGCAGGTTTGAAGGAA 4379
Qy 1352 er---ProSerLeuProAspGlnAspLysAlaGlyCysIleGluValGlnValGlnSerL 1371
Db 4380 GCCCTCCTCCCTGCTAGTCAAGGAGGAGGAGTATGCACCAAAATTCAGATTTCAGAGCT 4439
Qy 1371 euAspThrThrValThrGlnThrAlaGluAlaValGluLysValIle-----GluThrV 1389
Db 4440 CTAGGCGCATCTTCACTCTAACAGCGGCTGCAGAGGAGGAAAGGTTCTAGGAGAACTG 4499
Qy 1389 alValIleSerGluThrGlyGlySerProGluCysValGlyAlaHisLeuLeuProAlaG 1409
Db 4500 CCAACATTTTAGAAACAGGGAACAGTGTGGAGCTGCAGGTGCACATTTAGTTCTGGAAG 4559
Qy 1409 lulyserSerAlaThrGlyGlyHisThrThrLeuGlnHisAlaGluAspThrValProL 1429
Db 4560 AGAAATCCTCGAAAAAATGAAGACTTTGCCGCTCATCCAGGGGAGATGCTGTGCCCA 4619
Qy 1429 euGlyProGluSerGlnAlaGluSerIleProIleIleValThrProAlaProGluSert 1449

```

```

Db 4620 CAGGGCCGACCTGTCTAGGCAAAATCGACACCACTGATAGTATCTGCTACTACCAAGAAAG 4679
Qy 1449 hrLeuHisProAspLeuGlnGlyGluIleSerAlaSerGlnArgGluAArgSerGluGluG 1469
Db 4680 GCTTAAAGTTCCGACCTTGAAGAGGAGAGAAACACACATCACTGAAGTGAAGTCAGATGAAG 4739
Qy 1469 luAspLysProAspAlaGlyProAlaAspGlyLysGluSerThrAlaIleGluLysV 1489
Db 4740 TCGATGAGCAGGTTGCTTCCAGGAGGTC-----AAAGTGAGTTAGCAATTTAGGAGG 4793
Qy 1489 alLeuLysAlaGluProGluIleLeuGluLeuGluSerLysSerAsn-LysIleValLeu 1508
Db 4794 ATTAGAGCTGAAATATGGGATTTTGAACTTGAGACCAAAAGCAGTAAAACTTGTCCAA 4853
Qy 1509 AsnValIleGlnThrAlaValAspGln-PheAlaArgThr---GluThrAlaProGluTh 1527
Db 4854 AACATCATCCAGACACGCGTTGACCAAGTTTGTACGTACAGAAAGAACAGCCACCCAAAT 4913
Qy 1527 rHisAlaTyrrAspSerGlnThrGlnValProAlaCysArgLeuAspSerArgGluProAs 1547
Db 4914 GTTGACGTCTGAGTTACAGACAA-----GCTTCACGTGATAAAAGCTGACAGCCAGG 4967
Qy 1547 nArgCysTrpThrLysMetLysAspAlaLys-----Me 1558
Db 4968 ACCG---TGGACAGGAAACCGGAGAAAGAGGAGGAGACCTCAGCGCTCTGCACAGGAT 5024
Qy 1558 tLysHisProValProGlnProArgGlu-----AspLeuGlnVa 1571
Db 5025 GAAACACCAATTACTTCAGCCAAAGAGGAGTCAGAGTCAACCCAGTGGGACAGACAT 5084
Qy 1571 lLeuThrValLeuGluAlaTrpAlaGlnProArgLys 1583
Db 5085 TCTGATATTTCCAAAGACATGAGTGAAGCTCAGAAA 5121
RESULT 9
ABL23161
ID ABL23161 standard; DNA; 8307 BP.
XX
AC ABL23161;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 20956.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ds.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE ) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Claim 1; SEQ ID NO 20956; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and

```


Db 3382 GATCAGAGCACTGAAAGTGTGAAACGATGAGGAATCTACCGCCAAAGCCCGACAATAAG 3441
QY 482 ILeValSerGluValGluMetLeuSerSerGlnGluArgIleLys----- 496
Db 3442 ATTGACGAGGATGTTTCAGAGATCAGCACTGAGAGATTCCTCAAGGATGTCATCATGCT 3501
QY 497 -----ValGlnGlySerProLeuLysLysLeuPheSerSerSerGlyLeuLysLysLeu 514
Db 3502 ACTGGCATCTAGCAGCGCTATCTCATGTC----- 3534
QY 515 SerGlyLysLysGlnLysGlyLysArgGlyGlyGlyAspGluGluProGlyGluTyr 534
Db 3535 -----AAGCCCGATGAAGAG 3549
QY 535 GlnHisIleHisThrGluSerProGluSerAlaAspGluLysGlyGlySerSerAla 554
Db 3550 ATCCAACCACTGACCTCGGTTCTGCACAGTTTGATCAA-----TCTACGACAGCT 3600
QY 555 SerSerProGluGluProGluGluThrThrCysLeuGluLysGlyProLeuGluAlaPro 574
Db 3601 AAGGTTGATAAAGAACCAATTGATGAATCCGCTGAGGATGAAGAACCAATT----- 3651
QY 575 GlnAspGlyGluAlaGluGluGlyThrThrSerAspGlyGluLysLysArgGluGlyLe 594
Db 3652 -----GGCAATCCGAGGAGGATTGAAACCAATTGATGAATCTGAGGAGGAC----- 3699
QY 595 ThrProTrpAlaSerPheLysLysMetValThrProLysLysArgValArgArgProSer 614
Db 3700 -----AAGAAACCGTTGAAGAATCTGCAAGATAGAAACCCGTT 3741
QY 615 GluSerAspLysGluGluGluLeuLysValLysSerAlaThrLeuSerSerThrAsp 634
Db 3742 GAA---GATTCCGAAGAAGAGAGACCATTCGCGACTGTCACTCCAGCGAGTGAGATT 3798
QY 635 SerThrValSerGluMetGlnAspGluValThrValGlyGluGluGlnLysPro--- 653
Db 3799 GAAAAGAAATCCAGCCGAGGAGCAAGAAAGACTGAGCTGACTTTGGCGCGCCACT 3858
QY 654 GluGluProLysArgArgValAspThrSerValSerTrpGluAlaLeuLysCysValGly 673
Db 3859 GAACAGCCAGAGGTCAACTCCGCCAGATTGCCACACTGCTGAGAGAGGAGTTGAT 3918
QY 674 SerSerLysArgAlaArgLysAlaSerSerSerAspAspGluGlyGlyProArgThr 693
Db 3919 GATAAGCTGGCAACCATCTGCTCCGTTTCTGCCGAAGTAGCTCAAGCCT----- 3972
QY 694 LeuGlyGlyAspSerHisArgAlaGluGluAlaSer---LysAspLysGlu-----Ala 710
Db 3973 ---GCCGATGAAAGAAACGCCTGAGACAGCGCAGATCCAGATGCTGAAATCCCTGCT 4029
QY 711 GlyThrAspAlaValProAlaSerThrGln----- 720
Db 4030 AGCACCGATGCGCCGAGTCTCGACTGAGTGGCTGCTGATCTCGACAGAAACCA 4089
QY 721 GluGlnAspGlnAlaGlnGlySerSerSerProGlu-----Pro 733
Db 4090 GAGGAGGATTCGACTAAAGCACCAGCGCGCTGAGTCTGATAGGTTCTCGAAGTTCCG 4149
QY 734 AlaGlySerProSerGluGlyGlyValSerThrTrpGluSerPheLysArgLeuVal 753
Db 4150 ACCAGTCTTCTACTGAAATGAG---ATTGAGGAGTCTGACAAATTCACGACAGTCGCT 4206
QY 754 ThrProArgLysLysSerLysSerLysLeuGluLysAlaGluAspSerSerValGlu 773
Db 4207 CCACCAAAATTTCTGTTCTGACAAACAGAGCCCACTGCTGAGGAAGATCTCGTC--- 4263
QY 774 GlnLeuSerThrGluIleGluProSerArgGluGluSerTrpValSerIleLysLys--- 792
Db 4264 ---CCTGCAACCTTTGAGCCCATCGAAAGCGAATTTGAAGTTAGCACCAAGAACCCA 4317
QY 793 -----PheIleProGlyArgArgLysLysArg----- 801
Db 793 -----PheIleProGlyArgArgLysLysArg----- 801

Db 4318 GCCGTGCAAGGACCACCACTGCCACTCTTGTCTCGCTCAGCCAGAGAAGAACCGATC 4377
QY 802 -----AlaAsp---GlyLysGlnGluGlnAlaThrValGluAsp 813
Db 4378 GATGCAGAGACTTCCACTGAGGCTGATATCAGCACCAACCAAGCGCTGAAGTTGAAAG 4437
QY 814 SerGlyProValGluLeuAsnGluAspAspProAsnVal----- 826
Db 4438 GAAGCTTCTGGTGAACATCCGAGTCGGACAATGAATCGATGCTGGTCTCTAGTACT 4497
QY 827 ProAlaValValProLeuSerGluTyrAsnAlaValGluArgGluLys---MetGluAla 845
Db 4498 CTTGTTCTCTGCTGCGGATGAGGACAGACTCCAGAGCACTGAGAAAGCTGTCGAGGCC 4557
QY 846 -----GlnGlyAsnThrGluLeuPro 852
Db 4558 GACGACAAAGTTTCACTGTTGCGCCACTTGTGGTGATGAAGAGGAGTCCAACCTTCC 4617
QY 853 GlnLeuLeuGlyAlaValTyrValSerGluGluLeuSerLysThrLeuValHisThrVal 872
Db 4618 AAGCTACCAAGATATCTTT-----GAAGAGGAAGCGCTGTTGCGGTTTACCACCGCA 4671
QY 873 SerValAlaValIleAspGlyThrArgAlaValThrSerValGluGluArg----- 889
Db 4672 GCTCCATCAAGGACGACGCTGAGGACCGAGTGAAGTTGAAGAGAAACCGATCGAA 4731
QY 889 ----- 889
Db 4732 GATGGACAAAAGCCTATTGAGGACGAAACCTCTACACCTACTTCACTGAAAATGAGATT 4791
QY 890 -----SerProSerTrpIleSerAlaSerValThrGluProLeuGluHis 904
Db 4792 GAGCCTGAATCTGACCGTGCACTACTATTGCTCTTCTAAGGAAGAACCTTCGGAACCA 4851
QY 905 ThrAlaGly-----GluAlaMetProValGluGluValThrGluLys 919
Db 4852 TCACCTGGTCTCTACTAAGATGAACCTGCCGAACCACTCCACTGATGCTCTCTGATCT 4911
QY 920 AspIleIleAlaGluThrPro-----ValLeuThrGlnThrLeuProGluGly 936
Db 4912 GAT---GAGAGCAAGAAACGCTGAGTCGGAAGTGCCACAACTGTTGCACCTGCTGTT 4968
QY 937 LysAspAlaHisAspAspMetValThrSerGluValAspPheThrSerGluAlaValThr 956
Db 4969 GAGAAGATTCCAAACAAGCAGTACTCTCCGATCAGGAACCACTGCACTTCTCTCTCT 5028
QY 957 AlaThrCluThrSerGluAlaLeuArgThrGluGluValThrGlu----- 971
Db 5029 GTTGCCCAAGCCTGATGAAGATGATAGAAAGAAACCGACACCCGAAATTCCTGATG 5088
QY 972 ---AlaSerGlyAlaGlu-----ThrThrAspMetValSerAla----- 984
Db 5089 CTTGATCTGTCGAAGGATGAGAATTCCTCTACTGATCAAAATTCATCTGAGGTACCA 5148
QY 985 -----ValSerGlnLeuThr 989
Db 5149 GAAAAGAAACCTGAAACTCTGCACAGACGCCAGAAAGAGATATTGTTGGCGCACT 5208
QY 990 AspSerProAspThrThrGluGluAlaThrProValGlnGluVal----- 1004
Db 5209 GCAGCTCCAAACGACCTCTGATGAAGTGCGCCCGTCCAGAGGCTGCTGAGGAAGTCTCT 5268
QY 1005 -----GluSerGlyValLeuAspThrGluGlu----- 1013
Db 5269 GCTGAGATTCCGAGCCCTCGACTGAACTGGAAATTAAACAGCAAGCAAGCAACTGCT 5328
QY 1014 -----GluGluArgGlnThr 1018
Db 5329 GCTCCAAGTATTGACCGCAAGAACCTTATGTTACTGAAATCGATGAAGAAGCCACCA 5388
QY 1019 GlnAlaIleLeuGlnAlaValAlaAspLysValLysGluGluSerGlnValProAlaThr 1038
Db 5389 GTGGCTCTATCTCCGAGAGGATGAGAACCACTGAGGA----- 5430

QY 1039 GlnThrValGlnArgThrGlySerIysAlaLeuGlnValGlnGluValGlnGluAsp 1058
 Db 5431 -----GAGAAACCGATGGAAACCAAGCCAACTGGAAAGAGCCCA 5469
 QY 1059 SerGluValLeuAlaSerGlnLys-----GlnLysAspValMetProLysGlyProVal 1076
 Db 5470 TCCGAGAGAAAGAGAGAGAAACCCATTGAAACAGATGATCTACTAGAGGAGCCAGTA 5529
 QY 1077 Gln---GlnAlaGlyAlaGlnHisLeuAlaGlnGlySerGluThrGlyGlnAlaThrPro 1095
 Db 5530 TCTACCGAGAGCTTCTGAAGCTGGATCTACCGAATCATCTGAGAGATTAAAGCCCTGACT 5589
 QY 1096 GlnSerLeuGlnValProGluValThrAlaAspValAspHisValAlaThrCysGlnVal 1115
 Db 5590 GAGGGA---GAAATGGCCGAAAGCCTGAGCAAGCAGCTTCTCGACTCCAGGCT 5646
 QY 1116 IleLysLeuGlnGlnLeuMetGlnGlnAlaValAlaProGlnSerSerGluThrLeu--- 1134
 Db 5647 -----CCAGTTGAGACAAATTCGGAATCTCCAGTAAGTGGCA 5685
 QY 1135 ---ThrAspSerGluThrAsnGlySerThrProLeuAlaAspSerAspThrAlaAspGly 1153
 Db 5686 GCTCAAGATGGGACAAAGCCCAAGCCAGGCTCCTGTTGACTCCGATGAAGATACCTCC 5745
 QY 1154 ThrGlnAspGluThrIleAspSer-----GlnAspSerLysAla 1167
 Db 5746 GCTCGTCGCGAGAGAGATTCCTCGTTCTGCGAGAGAGTGGAGGCGCCAGAAAGTC 5805
 QY 1168 ThrAlaAlaValArgGlnSerGlnValThrGlnGlnGlnAlaAlaThrAlaGlnLysGlu 1187
 Db 5806 ACACTGCTCTCACTCAAGCT---GCCGAGAGAGACGAACTTAAAGACAGCCCGGAGACT 5862
 QY 1188 GlnProSerThrLeuProAsnAsnValProAlaGlnGlnGlnHisGlyGlnGluProGly 1207
 Db 5863 GAACCATCTTCC---ACTGACAGAGTACCAAGAACTGAATACAGAGAGCCGAG--- 5913
 QY 1208 ArgAspValLeuGlnProThrGlnGlnGlnLeuThrAlaAlaAlaValProValLeuAla 1227
 Db 5914 -----GACGAGACCAAGGCTGACGAGACCCCTGAGTGTGTACCCCAAGTGTCT 5961
 QY 1228 LysThrGlnValGlyGlnGlnGlnGlnValAspTyrPheLysAspGlyGlnLys----- 1244
 Db 5962 GATGTGGCAACATCCACTTCGCGCCCACTGCTGTGTGATGATGAAAAGAGCAACA 6021
 QY 1245 -----ValLysGlnGlnGlnGlnValPheValHisSerGlyProAsnSer 1259
 Db 6022 GCCACCACTGCTCGCTGCAAGAGAGAGAGAGATC-----AAGCCAACTATC 6069
 QY 1260 GlnLysAlaAlaAspValThrTyrAspSerGlnValMetGlyValAlaGlyCysGlnGlu 1279
 Db 6070 GCTCAGCTGCTGAAATTCCTCAACCATC-----GAA 6102
 QY 1280 LysGlnSerThrGlnValGlnSerLeuSerLeuGlnGlnGlnGlnMetGlnThrAspVal 1299
 Db 6103 AAGAGCCCGGTGATGAGCAA-----GAAATGGAATCCGAAACCAAGGCCACTCCAGCA 6156
 QY 1300 GlnLysGlnLysArg-----GluThrLysProGlnGlnValSerGlnGlnGln 1316
 Db 6157 GAGTCTGACGGCCAGCAATTGATGAAATGGCCCTGCAACATCTGAGCCCATCATGAG 6216
 QY 1317 GlnGlnThrAlaAlaProGlnHisGlnGly----- 1326
 Db 6217 GCTTCACAGCGCGCTCCCAAAAGAGGAAATCCACAACGTTGTAGCGCAGCTTGCT 6276
 QY 1327 -----ThrTyrGlyLysProValLeu----- 1333
 Db 6277 GCTGTTCAAGATGAATTAAGAGATGTCAACCACTCAACCAAGATTGCGCATGAGAG 6336
 QY 1334 ThrLeuAspMetProSerSerGlnArgLysLysAlaLeuGlySerLeuGlyLysPro 1353
 Db 6337 GAAATGGCTCTCTCAAGAGCAAAAGAAACATCATTCATGATGTTTCCACCGATTCACCC 6396

QY 1354 SerLeuProAspGlnAspLysAlaGlyCysIleGlnValGlnValGlnSerLeuAspThr 1373
 Db 6397 ACTGCCCAAGATGATGAGAAACAAAGCAAAACGAGGCTCCAGTG-----GCACCCACC 6450
 QY 1374 ThrVal-----ThrGlnThrAlaGlnAla 1381
 Db 6451 ACTGTTTATTCGCCGAGCTGCTGATTTCTCTGCTGATTTACTTACTCCACAGTGAAGTA 6510
 QY 1382 ValGlnVal-----IleGluThrVal-----ValIleSerGluThrGlyGlu 1396
 Db 6511 CCATCTCCTGTTGAATTCACACGAAAGCCAAATGATGATATCATCTCCACACTATTCCT 6570
 QY 1397 SerProGlnCysValGlyAlaHisLeuLeuProAlaGlnLysSerSerAlaThrGlyGly 1416
 Db 6571 CCTCATCCGCGGATGTGTGCTGCGCTCCACGTCCACTGAAGTGAAGACAGAGCTCCGCTT 6630
 QY 1417 HistThrLeuGlnHisAlaGluAspThrValProLeuGlyProGlnSerGln----- 1434
 Db 6631 ACTGTGTCTCTCAGATGATCTGAAAAGACT---CCAGTTTCCCGGCTCCCGAGATTCT 6687
 QY 1435 -----AlaGlnSerIleProIleIleValThr 1443
 Db 6688 GATAAGATCCGAGCTCAGAGCTCCCGAGGATGCCGAAATCCAGTAACAGCCACA 6747
 QY 1444 ProAlaProGlu-----SerThrLeuHisProAspLeuGlnGlyGlnLysSer 1459
 Db 6748 CCTTTGACGACCAACAGATCCAGCCACTGTGCTCTCTCAGACTGATGATGGGTTTCA 6807
 QY 1460 AlaSerGlnArgLysArgSerGlnLysAspLysProAspAlaGlyPro---AspAla 1478
 Db 6808 GCTACCGAGCTCTCTTGTGATGAGACAAAGATCCAAATCTGCGGCTCTCTTGATAG 6867
 QY 1479 AspGlyLysGlnSerThrAlaIle-----GlnLysVal---LeuLysAlaGlu 1493
 Db 6868 GAGAAAGATTCATCCACGAGCTCTCTTGATGATGAAAGATTCCTGCCAGTTAGC 6927
 QY 1494 ProGlnIleLeuGlnLysSerLysSerAsnLysIleValLeuAsnValIleGlnThr 1513
 Db 6928 CCGGTGTTTTCGATGTGAGACCAAGCAGCGAGAAG-----CCT 6966
 QY 1514 AlaValAspGlnPheAlaArgThrGlu---ThrAlaProGluThrHisAlaTyrAspSer 1532
 Db 6967 GCTGATCTGAGTATGATGTGAAGAGAGACCGAACCACTGTCATGATGTGGAGACC 7026
 QY 1533 GlnThrGlnValProAlaCysArgLysAspSerArgLysProAsnArgCysTyrThrLys 1552
 Db 7027 AGCACTGATGAGCTACT----- 7044
 QY 1553 MetLysAspAlaLysMetLysHisProValProGlnProArgGlnLysLeuValLeu 1572
 Db 7045 ---TCTGATGCCAAACTGAAACCACTTCTGCACTGCTAGCGCATTCGAA---TCT 7098
 QY 1573 ThrValLeuGlnAlaTyrAlaGlnProArgLysCysLeuProArgLeuGlnLysAla 1592
 Db 7099 CCAGCCACGAGGCGGAGATTTGCTCCAGAACTGCAGGCCCAAGACTTGAAAGAGGCTG 7158
 QY 1593 Pro 1593
 Db 7159 CCA 7161
 RESULT 10
 ABL23160
 ID ABL23160 standard; DNA, 11597 BP.
 XX ABL23160;
 XX 26-MAR-2002 (first entry)
 DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 20953.
 XX Drosophila; developmental biology; cell signalling; insecticide;
 KM pharmaceutical; gene; ds.
 XX

OS Drosophila melanogaster.
 PN WO200171042-A2.
 XX PD 27-SEP-2001.
 XX PF 23-MAR-2001; 2001WO-US09231.
 XX PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX (PEKE) PE CORP NY.
 PA Venter JC, Adams M, Li PWD, Myers EW;
 PI WPI; 2001-656860/75.
 XX
 DR New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Claim 1; SEQ ID NO 20953; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABBS7737-ABBS72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 11597 BP; 3279 A; 3082 C; 2895 G; 2341 T; 0 other;
 Alignment Scores:
 Pred. No.: 1,29e-23 Length: 11597
 Score: 639.50 Matches: 389
 Percent Similarity: 34.82% Conservative: 227
 Best Local Similarity: 21.99% Mismatches: 707
 Query Match: 7.92% Indels: 447
 DB: 23 Gaps: 69
 US-09-902-432-4 (1-1596) x ABL23160 (1-11597)
 QY 3 AlaGlySerSerThrGluGlnArgSerProGluGlnProAlaGlySerAspThrProSer 22
 DB 4870 GCAAAGACCACCAACCGAGCAGCCCAAGGAGGAGTCTCTACTGAGGCTGAGGACGCT 4929
 QY 23 GluLeuValLeuSerGlyHisGlyProAlaAlaGluAlaSerGlyAlaAlaGlyAspPro 42
 DB 4930 GAATCGAAGTTACCTAGTAGTCCCGCATAAACAGAGGTGCCAGAGCTGAACCA 4989
 QY 43 AlaAspAlaAsp-----ProAlaThrLysLeuProGlnLys 54
 DB 4990 GCAGACAAGGACCATAAGGATGAGGAAGATGTCACAGACCGCCACTGATTTGCCAATCAA 5049
 QY 55 AsnGly-----GlnLeuSerSerValAsnGlyValAlaGluGlnGly 68
 DB 5050 TCTGATATTGGCCACCCTGTTGTGCATACCGAAGCTACTACGGGTGACGCTGAGACCACT 5109
 QY 69 Asp-----ValHisValGln-----GluGluAsn 76
 DB 5110 GATGAGACTGCTACCGTAAGCCACCATCTGTTATCTGCCTCTGTTGGAGAGAAGGAT 5169
 QY 77 GlnGluGlyGlnGluGluValValAspGluAspValGlyGlnArgGluSerGluAsp 96
 DB 5170 ACTGAGAAAGAGGACGAACCTACTGCTACAGAGCCATCTGTCCAAGAGAACCATCTAAG 5229
 QY 97 ValArgGluLysAspArgValGluGluMetAlaAlaAsnSerThrAlaValGluAspIle 116
 DB 97 ValArgGluLysAspArgValGluGluMetAlaAlaAsnSerThrAlaValGluAspIle 116

DB 5230 CTTGTGGATATCTCTCCAGTGTCCAGGAAGTTCCGTCCAGCACTGCTAAAGTTGACAAT 5289
 QY 117 ThrLysAsp-----GlyGlnGluGluThrSer 125
 DB 5290 CGCAACGACTTCGAAACGGAGAAACCAACCTGCCACCAAGTGGCGAGGATCAGAGCTCC 5349
 QY 126 GluIleIleGlu-----GlnIleProAlaSerGluAsnAsnValGluGluMetValGln 143
 DB 5350 GAGCCCTGCCCGCATGATCTGCCCGCTGGTATTCTCTGGCGAAGGTGACTGCTGGTG 5409
 QY 144 ProAlaGluSerGlnAlaAsn-----AspValGly 153
 DB 5410 GAGGGCAAGACCTATGCCAACACCAACCATCTCTCCCGCTACCGCACCTCGCGAGCTCTCC 5469
 QY 154 PheLysLysValPheLysPheValGlyPhe----- 163
 DB 5470 TGCAAAATGTTATCAGCACCTCGTGGCTGCAACAGATGAGTGTCAAGTTGCCCGAGAAC 5529
 QY 164 -----LysPheThrValLysLysAspLysAsnGluLysSerAspThrValGlnLeuLeu 181
 DB 5530 TTGGAGAAATGTATCCGTGGCTGCCGATCTTCTGGATGGTTGTCTCTCCACTTACATTTCG 5589
 QY 182 ThrValLysLysAspGluGlyGluGlyAlaGluAlaSerValGlyAlaGlyAspHisGln 201
 DB 5590 GATGAGAGCACTGAAGTGTCTGAAAGGATGAGAAATCTACCCCAAGCCCGACAATAAG 5649
 QY 202 GluProSerValGluThrAlaValGlyGluSerAlaSerLysGlu----- 216
 DB 5650 -----ATTGACGAGGATGTTTCAGAGATCAGCACTGAAGAGATTCACCAAGATGTC 5700
 QY 217 -----SerGluLeuLysGlnSerThrGluLysGlnGluGluThrLeu 230
 DB 5701 ATCATGCTACTGGCATCACTGAGCAGCTCTATCTCATGTCAGGCCGAT----- 5751
 QY 231 LysGlnGluGlnSerSerThrGluIleProLeuGlnAlaGluSerAspGlnAlaAlaGlu 250
 DB 5752 GAGAGATCCCAACAGTGACCTCGGTTCTCGACAGTTTGTGAATCATAGACAGCTAAG 5811
 QY 251 GluGluAlaLysAspGluGlyGluGluLysGlnGluLysGluProThrLysSerProGlu 270
 DB 5812 GTTGAATAAGAAACCAATTGATGAATCCGCTGAGATAAGAAACCAATTTGGCAATCCGAG 5871
 QY 271 SerProSerSerProValAsnSerGluThrThrSerSerPheLysLysPhePheThrHis 290
 DB 5872 GAGGATTCGAAACCAATTGAT----- 5892
 QY 291 GlyTrpAlaGlyTrpArgLysLysThrSerPheLysLysSerLysGluAspAspLeuGlu 310
 DB 5893 -----GAATCTGAGGAGGACCAAGAA 5913
 QY 311 ThrAlaGluLysArgLysGluGlnGluAlaGluLysValAspGluGluLysGluLys 330
 DB 5914 CCGGTTGAAGAATCTCAGAAAGATAGAAACCCGTTGAAGATTCGGAAGAGAGAGAGAG 5973
 QY 331 -----ThrGluProAlaSerGluGluGlnGluProAlaGluAspThrAspGln 346
 DB 5974 CCATTGCGGACTGTCATCCAGCGAGTGAGATTGAAAGAAAGATCCAAAGCCCGAGGACGAA 6033
 QY 347 AlaArgLeuSerAlaAspTyr-----GluLysValGluLeuProLeuGluAsp 362
 DB 6034 AAGAAGACTGAGGTGACTTTTGGCGGCCCACTGAACAGCAGCAGGCTACAACTCCCGCC 6093
 QY 363 GlnValGlyAsp--LeuGluAlaSerSerGluGluLysCysAlaProLeuAlaThrGlu 381
 DB 6094 CAGATTCGCCGACTGCTGAGAGGAGGTTGATGAAGCTGGCAACCATCTGTCTCC 6153
 QY 382 ValPheAspGlu-----LysMetGluAlaHisGln 391
 DB 6154 GTTCTCGCGAAGATGAGCTCAAGCTCGCGATGAAAGAAACCGCATGAGACAGCGCAG 6213
 QY 392 GluValValAlaGluValHisValSerThrValGlu----- 403
 DB 6214 ATCCAGATGCTGAAATCCCTGTGTAGCACCGCATGAGCCGAGTCTCTCGACTGAGTGCCT 6273

QY 404 -----LysThrGluGluGluGluGlyGlyGlyGluAlaGluGly 417
 DB 6274 ACTGTGATCTCGACAGAACCCAGGAGATCGACTTAAGACCCAGCG----- 6327
 QY 418 GlyValValValGluGlyThrGlyIleuSerLeuProProGluLysLeuAlaGluProGln 437
 DB 6328 -----CTGAGTCTGATTAAGTTCTT----- 6348
 QY 438 GluValProGlnGluAlaGluProAlaGluIleuMetLysSerArgGlyMetCys--- 456
 DB 6349 GAAGTTCGACAGAGCTTCTACTGAAATAGATGAGATCGACCAATTCACAGACA 6408
 QY 457 -----ValSerGlyGlySerHisThrGlnLeuThrAspLeuSerProGlu 471
 DB 6409 GTGCTCCACCAAAATTTCTCTCTGACGAAACAGACCCT-----GCTGAG 6459
 QY 472 GluLysThrLeuProLysHisProGluGlyIleValSerGluValGluMetLeuSerSer 491
 DB 6460 GAAGATCTGCTCCGCAACCTTTGAGCCCATCGAAAGCGAATTTGAGATT-----AGC 6513
 QY 492 GlnGluArgIleLysValGlnGlySerProLeuLysLysLeuPheSerSerSerGlyLeu 511
 DB 6514 ACCAAGAAAGCCGCGTGAAAGACACACACCTGCCACTTGGCTCGGCTCAGCCAGAG 6573
 QY 512 LysLysLeuSerGlyLysLysGlnLysGlyLysArgGlyGlyLysAlaGluGluPro 531
 DB 6574 AAGAAA-----CCAGTCGATCGAGACT 6597
 QY 532 GluLysLysGlnHisIleHisThrGlu---SerProGluSerAlaAspGluGlnLysGly 550
 DB 6598 TTCACGAGGCTGATATTCAGCCAGACGAGCGCTGAAGTTGAAAGAGAGACTTGTGT 6657
 QY 551 GluSerSerAlaSerSerProGluGluProGluGluThrThrCysLeuGluLysGlyPro 570
 DB 6658 GAACATCCGAGTCCGACAAATGA-----ATCATGCTGCTGCT 6696
 QY 571 LeuGluAlaProGlnAspGlyGluAlaGluGluGlyThrThrSerAspGlyLysLys 590
 DB 6697 TCTAGTACTCCTGTTCTGTTCTGCGATGAGACAAAGCTCCAAAGCACTGAAAGACT 6756
 QY 591 ArgGluGly-----IleThrProThrAlaSerPheLysMetVal 604
 DB 6757 GTCGAGGCGGACGACAACTTCCACTGTGGCCACTTGTGCTGATGAAAG----- 6810
 QY 605 ThrProLysLysArgValArgArgProSerGluSerAspLysGluGluLysLeuGlnLys 624
 DB 6811 ---GAGTCCAACTTGGCCCAAGCTACCAAGATATCTTGAAGAGAGAGCGCTGTGG 6867
 QY 625 ValLysSerAlaThrLeuSerSerThrAspSerThrValSerGluMetGlnAspGluVal 644
 DB 6868 GTTACCACCGGCTCCATCAAGGACGAGCTGACAGAACCCAGTCCGAAATTTGAAGAG 6927
 QY 645 LysThrValGlyGluGluGlnLysProGluGluProLysArgArgValAspHisSerVal 664
 DB 6928 AAACGATGAGAGATGAGAAAAGCTATTTAGAGCAAACTTACACCTTCTATCT 6987
 QY 665 SerThrGluAlaLeuIleCysValGlySerSerLysLysArgAlaArgLysAlaSerSer 684
 DB 6988 GAAATATGAG-----ATTAGCCTGATCTGACCGTGCAACTATTTGTCTCT 7035
 QY 685 SerAspAspGluGlyLysProArgThrLeuGlyLys----- 696
 DB 7036 TCTTAGGAAGAACTTCCGAAACCATCCACTGGTGTCTCTACTAAGATGAACTCCGAA 7095
 QY 697 AspSerHisArgAlaGluGluAlaSerLysAspLysGluAlaGlyThrAspAlaValPro 716
 DB 7096 CCATCACTGATGCTCTCTGATGAGACGAAAGAAAGCCGTGAGTCCGAAAGTGGCC 7155
 QY 717 AlaSerThrGlnGluGlnAspGlnAlaGlnGlySerSerSerProGluProAlaGlySer 736
 DB 7156 ACAACTGTTCGACCTGCTGAGAAAGATTCCAAACAGAGATATCTCCGATGAGGAA 7215

QY 737 ProSerGluGlyGluGlyValSerThrTrpGluSerPheLysArgLeuValThrProArg 756
 DB 7216 CCA-----ACTGCCACT 7227
 QY 757 LysLysSerLysLysLeuGluGluLysAlaGluAspSerSerValGluGlnLeuSer 776
 DB 7228 TCTGCTCTGTTCGCAACCCGATGATGATGAGAAAGAAACAGACCGAAATTTCC 7287
 QY 777 ThrGlnIleGluProSerArgGluGluSerTrpValSerIleLysLysPheIleProGly 796
 DB 7288 ACTGACGCTCTGCAATCTCCGAAAGAG----- 7314
 QY 797 ArgArgLysLysArgAlaAspGlyLysGlnGluGlnAlaThrValGluAspSerGlyPro 816
 DB 7315 -----GATAGAAATTTCTCTACTGATTCAAATTCCA 7344
 QY 817 ValGluIleAsnGluAspAspProAsnValProAlaValAlaProLeuSerGluLysAsn 836
 DB 7345 TCTGAGTACCAAGAAAGAAACCTGAACTCTGCAAGAGCCCA----- 7389
 QY 837 AlaValGluArgGluLysMetGluAlaGlnLysThrGluLeuProGlnLeuGlnLys 856
 DB 7390 -----GAAAGAGAGATTTGTGTGGCGCACTGCAGAGCTCCAAAGCCTGTGAT 7437
 QY 857 AlaValIlyrValSerGluGluLeuSerLysThrLeuValHisThrVal----- 872
 DB 7438 GAAGTCCCCCGGTCCAGAGGCTGCTGAGAAAGCTCTGTGATTCGCCAGCCCTCG 7497
 QY 873 -----SerAlaAlaValIleAspGlyThrArgAlaValThrSerValGluGluArg 889
 DB 7498 ACTGAAACTGGAATTAAACAGACAGAAACAACTGCTGCTCCAAAGATTTGACCGCAAG 7557
 QY 890 SerProSerTrpLysSerAlaSerValThrGluProLeuGlnHisThrAlaGlyGluAla 909
 DB 7558 GAACCC-----TATGTACTGAAATCGATGAAGAACCCACC-----ACA 7596
 QY 910 MetProProValGluGluValThrGluLysAspIleIleAlaGluGluThrProValLeu 929
 DB 7597 GTGGCTCTATCTCCGAAAGATGAGAACCA---ACTGAGAAAGAAACCACTG--- 7650
 QY 930 ThrGlnThrLeuProGluGlyLysAspAlaHisAspAspMetValThrSer----- 946
 DB 7651 ---GAACGAAGACCACTGAGAAAGACCATCCAGAAAGAAAGAAAGAAACCATTT 7707
 QY 947 GluValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeuArgThr 966
 DB 7708 GAACAAGATGATCTACTGAGGAGACAGATATCTACCGAAAGCTTGAAGCTGAGATCTAC 7767
 QY 967 GluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSerAlaValSer 986
 DB 7768 GAA----- 7770
 QY 987 GlnLeuThrAspSerProAspThrThrGlnGluAlaThrProValGlnGluValGluSer 1006
 DB 7771 -----TCATCTAGAAAGTTAAGCCCTGCACCTGAG----- 7800
 QY 1007 GlyValLeuAspThrGluGluGlnGluArgGlnThrGlnAlaIleLeuGlnAlaValAla 1026
 DB 7801 GGAGAAAGTTGCCGAAAGACCTGAGACAGACCTTCTGACGTGCCAAGCTCCAGTT 7860
 QY 1027 AspLysValLysGlu---GluSerGlnValProAlaThrGlnThrValGlnArgThrGly 1045
 DB 7861 GAGACAAATTCGGAATCTCCACTGAATGCGACGTCAAGATGGCAGAACGCCCAAGC 7920
 QY 1046 SerLysAlaLeuGluLysValGluGluValGlnGluAspSerGluValLeuAlaSerGlu 1065
 DB 7921 GAGGCTCTGTGAGCTC-----GATGAAGATACCTCGCTCCGCTCCAGCAGAG 7968
 QY 1066 LysGluLysAspValMetProLysGlyProValGlnGluAlaGlyAlaGluHisLeuAla 1085
 DB 7969 AAGATTCCCTCGTT-----TCTGGGAGAGAA-----GTC 7998
 QY 1086 GlnGlySerGluThrGlyGlnAlaThrProGluSerLeuGluValProGluVal----- 1103

[illegible]

Db	8874	TTCCCCGGCTCCCGCAGGATTCTGATTAAGACTCCGAGCTCAGAGGCTCCCGCAGGATGCCGA	8993
Qy	1436	userIleProIleIleValThrProAlaProGlu-----SerThrLeuHisPr	1452
Db	8934	CGAAATCCGAGCTACACGCACCTTTTCAGACACACAGATCCCGAGCCTCTGGGTCC	8993
Qy	1452	oAspLeuGlnGlycyluIleSerAlaSerGlnArgGluArgSerGluGluAspLysPr	1472
Db	8994	TCAGACTGATGATGGGTTCAGCTACCGCAGCTCTCTTGATGAAGACAAGATCCAAAC	9053
Qy	1472	oAspAlaGlyPro--AspAlaAspGlyLysGluSerThrAlaIle-----G	1487
Db	9054	TACTGCGGCTCCTTTGATGAGGAGAAGATCCATCCACAGCAGCTCCTTTGATGATGA	9113
Qy	1487	uLysVal--LeuLysAlaGluProGluIleLeuGluLeuGluSerLysSerAsnLysIl	1506
Db	9114	AAAGATTCTGCTCCAGTTAGCCCGCTGTTTTCGATGTGGAGCCACACGAGAGAG--	9171
Qy	1506	eValLeuAsnValIleGlnThrAlaValAspGlnPheAlaArgThrGlu---ThrAlaPr	1525
Db	9172	oGluThrHisAlaTyRAspSerGlnThrGlnValProAlaCysArgLeuAspSerArG	1545
Qy	1525	oGluThrHisAlaTyRAspSerGlnThrGlnValProAlaCysArgLeuAspSerArG	1545
Db	9213	ACCTGTCATGATGTGGAGCAGCAGCTGATGAGCCTACT-----	9252
Qy	1545	uProAsnArgCysTrpThrLysMetLysAspAlaLysMetLysHisProValProGlnPr	1565
Db	9253	-----TCTGATGCCAACTGAAGCCACCTACTTCTTGCACC	9287
Qy	1565	oArgGluAspLeuGlnValLeuThrValLeuGluAlaTrpAlaGlnProArgLysCysLe	1585
Db	9288	TGCTACGCCATCGAA--TCTCCAGCCACGAGGCGGAGATGTGCCAGAACTGCAGC	9344
Qy	1585	uProArgLeuGlnLeuLysAlaPro	1593
Db	9345	CCCAGACTTTGAAAAGGAGGTGCCA	9369
RESULT	11		
ABL05592			
ID	ABL05592	standard; cDNA; 18506 BP.	
AC	ABL05592;		
XX			
DT	26-MAR-2002	(first entry)	
XX			
DE	Drosophila melanogaster	expressed polynucleotide SEQ ID NO 11258.	
XX			
KW	Drosophila; developmental biology; cell signalling; insecticide;		
KW	pharmaceutical; gene; ss.		
XX			
OS	Drosophila melanogaster.		
XX			
FN	WO200171042-A2.		
PD	27-SEP-2001.		
XX			
PF	23-MAR-2001; 2001WO-US09231.		
XX			
PR	23-MAR-2000; 2000US-191637P.		
PR	11-JUL-2000; 2000US-0614150.		
XX			
PA	(PEKE) PE CORP NY.		
XX			
PI	Venter JC, Adams M, Li PWD, Myers EW;		
XX			
DR	WPI; 2001-656060/75.		
DR	P-PSDB; ABB61489.		
XX			
PT	New isolated nucleic acid detection reagent for detecting 1000 or more		
PT	genes from Drosophila and for elucidating cell signalling and cell-cell-		
PT	interactions -		
XX			

PS Claim 1; SEQ ID NO 11258; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
CC sequences (AB101840-AB16175) and the encoded proteins
CC (AB161737-AB162072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIP
CC at ftp.wipo.int/pub/published_pcf_sequences.

XX Sequence 18506 BP; 5394 A; 4676 C; 4890 G; 3546 T; 0 other;

Alignment Scores:

Pred. No.:	1,366-19	Length:	18506
Score:	565.50	Matches:	412
Percent Similarity:	36.99%	Conservative:	282
Best Local Similarity:	21.96%	Mismatches:	737
Query Match:	7.00%	Indels:	448
DB:	23	Gaps:	81

US-09-902-432-4 (1-1596) x ABL05592 (1-18506)

QY 6 SerThrGluGlnArgSerProGluGlnProAlaGlySer----- 18
DB 10770 TCACAGAGAAAGCTCCCGGACAGATCTGAGCGAAGAGCTTGAAGATCGCGACGACCG 10829
QY 19 -----AspThrProSerGluLeuValLeuSerGlyHisGlyPro 31
DB 10830 TCGCAGAGACCTCTGCTCCGATCTGTACGAAAGCGTTAAGATGAAATCCCA 10889
QY 32 AlaAla-----GluAlaSerGlyAlaAlaGly-----AspProAla 43
DB 10890 GTCCGATCCAAAGAGACGCTCCAGACGATCCGTTGCGAAGAAATCCCAAGATTCGGCT 10949
QY 44 AspAlaAspProAlaThrLys-----LeuProGln---LysAsnGlyGlnLeuSer 59
DB 10950 GATGATTCAAAGAGACGACGACGATGCTGCTCAAGCGAAGCGGAGATATTAA 11009
QY 60 SerValAsnGlyValAlaGluGlnGlyAspValHisValGlnGluGlnGlnGlnGly 79
DB 11010 GAT-----GAAATATCACTTTGGCTTAAAGATGAAGCTGAAGTCT 11054
QY 80 GlnGluGlnGlu-----ValValAspGluAspVal 89
DB 11055 AAGAGAGAAATCAAGCGACAGTGGTGGCCGAACATTCCTTTGCTCCAGAGAAATT 11114
QY 90 GlnGlnArgGluSer-----GluAspValArgGlnLysAspArgValGln 104
DB 11115 TCCAGACCGACATCCGTTGACGAAAGTGTTAAGACGAAAGCTGAAGATTAAGAA 11174
QY 105 -----GluMetAlaAlaAsnSerThrAlaValGlnAspIleThrLysAsp 119
DB 11175 TCACCTTAAATGTCTTAAGAGAGCTTCACAGCAGATCCGTTCCAGAGAGTTAAGAA 11234
QY 120 GlyGlnGluGlnThrSerGlu-----IleIleGlnIleProAla 133
DB 11235 GAAGCTGAGAAAGTCCAAAGAGAAATCTAGGCGAGAGTCACTGTAATAATACCTTTG 11294
QY 134 SerGluAsnAsnValGlnGluMetValGlnProAlaGluSerGlnAlaAsnAspValGly 153
DB 11295 CCTCCAAAGAGAGCTTCAGAGACGAGCTCCGTTCCAGAGAAAGT----- 11336
QY 154 PheLysLysValPheLysPheValGlyPheLysPheThrValLysLysAspLysAsnGln 173
DB 11337 -----GTAAGAGATGAAGCTGAT 11354
QY 174 LysSerAspThrValGlnLeuLeuThrValLysLysAspGlnGly-----GlnGly 190
DB 11355 AAGTCC-----AAGAGAGAAATCTAGGCGAGAGTGGGG 11387

QY 191 AlaGluAlaSerValGlyAlaGlyAspHisGlnGluProSerValGluThrAlaValGly 210
DB 11388 GCGGAAATCT-----CCTTTGGCTCCAGAAAGCTTCC 11423
QY 211 GluSerAlaSerLysGluSerGluLeuLysGlnSerThrGlnLysGlnGlnGlyThrLeu 230
DB 11424 AGACGAGCTTCCTGTCGAGAAAGATTTAAGACGAGCTGAGAAAGTCTAAGAGAAATCT 11483
QY 231 LysGlnGlnGlnSerSerThrGlnIleProLeu-----GlnAlaGluSerAspGln 247
DB 11484 AGCGAGAGTGGTGTGCGTGAATAATCTCTTGGCTCCAGAGAGCTTCAGACCAACT 11543
QY 248 AlaAlaGlnGlnGlnAlaLysAspGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 267
DB 11544 TCCGTTGCAAGAAAGTGTAAAGATGAAGCT--GAGAAAGTCTAAGAGAAATCTAAGAA 11600
QY 268 SerProGluSerProSerProValAsnSerGluThrThrSerPheLysPhe 287
DB 11601 GACTCGGTGGCGGAAATCTCTTGGCTCCAGAGAGCTTC----- 11645
QY 288 PheThrHisGlyTyrPalaGlyTyrPalaGlyLysThrSerPheLysSerLysGluAsp 307
DB 11646 -----AGACGAGCTTCGTTGACGAAAGTGTTCAGGAC 11678
QY 308 AspLeuGluThrAla-----GlnLysArgLysGlnGlnGlnGlnGlnGlnGlnGlnGln 325
DB 11679 GAAGCTGAGAGAGTCTAGAGAAATCTAGGAGAGTGGTGGCGGAAATCTCTTGTG 11738
QY 326 GlnGlnLysGlnLysThrGlnProAlaSerGlnGlnGlnGlnGlnGlnGlnGlnGln 345
DB 11739 GCCTCCAGAGAGCTTCAGACGAGCTTCGTTGCGAAGAAATTAAGACGAGCTGAG 11798
QY 346 GlnAlaArg-----LeuSer 350
DB 11799 AAGCTTAAGAGAAATCTAGGCGAGAGTGGTGGCGGAAATCTCTTGGCTCCAA 11858
QY 351 AlaAspTyrGlnLysValGlnLeuProLeuGlnAspGlnValGlyAsp---LeuGlnAla 369
DB 11859 GGAGCTTCGAG--ACCCAACTTCCTGTCGCA--GAAAGTCTAAGAGAGAGCTGAGAG 11914
QY 370 SerSerGlnGlnLysCys-----AlaProLeuAlaThrGlnVal 382
DB 11915 TCTAAGAGAAATCTAGTCCGAGCTCGGTGGCGGAAATCTCTTGGCTGCTC----- 11968
QY 383 PheAspGlnLysMetGlnAlaHisGlnGlnValValAlaGlnValHisValSerThrVal 402
DB 11969 -----AAGAGAGCTTCAGACCAAGCTTCGTTGCGAAGAGTTCAGAGCGAAGCT 12019
QY 403 GlnLysThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 422
DB 12020 GAGAGAGTCTAAGAGAAATCTAGGCGAGAGTGGTGGCGGAAATCTCTTGTG----- 12073
QY 423 GlyThrGlnGluSerLeuProProGlnLysLeuAlaGlnProGlnGlnValProGlnGln 442
DB 12074 GCTCCAGAGAGCTTCAGACCAAGCTTCGTTGCGAAG-----AGCTTAAGAGATGA 12127
QY 443 AlaGlnProAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 462
DB 12128 GCTGAGAGTCCAAAGAA-----GATCTAGGCGAGAGTGGTGGCT----- 12169
QY 463 ThrGlnLeuThrAspLeuSerProGlnGlnLysThrLeuProLysHis---ProGlnGly 481
DB 12170 ---GAAATATCTCTTGGCTTCAGAGAGTTCAGAGAGCTTCGTTGCGTTCAGAGAAAGT 12226
QY 482 IleValSerGlnValGlnMetLeuSerSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 499
DB 12227 GTTAAGAGTGAAGCTGAGAGCTTAAGAGAAATCTAGGCG--AGAGTGGTGGCGGAA 12285
QY 500 -----SerProLeuLysLysLeuPheSerSerSerGlnLysLeuLysLeuSerGly 516
DB 12286 AATCTCTTGGCTTCAGAGAGAGCTTCAGACCAAGCTTCGTTGCGAAGAAAGTGTAAAG 12345

Qy	517	LysLysGlnLysGlyArgGlyGlyGlyAspGluGluProGlyGluThrGlnHis	536
Db	12346	ATGAAGCTCAGAAGTCTAAGGAAGAAATCAAGACGCGA-GAGTCGGTGCCGCCAAAATCTTCT	12404
Qy	537	IleHisThrGlusr-----ProGluSerAlaAspGluGlnLysGlyGluSerSer	553
Db	12405	TGGCGCTCCAGAAAGCTTCCAGACCAGCTTCCGTTCGAGAAAGTGTTAAGATGAAGACT	12464
Qy	554	AlaSerSerProGluGluProGluGluThrThrCysLeuGluLysGlyProLeu-----	571
Db	12465	GAGAAGTCTAAGGAAGAACTTAGCGCAGAGTCGGTGCCGCCGAAAAATCTCTTTGGCCCTCC	12524
Qy	572	---GluAlaProGlnAspGlyGluAlaGluGluGlyThrThrSerAspGlyGluLysLys	590
Db	12525	AAGGAAGCTTCCAGACCAAGCTTCCGTTCGAGAAAGTGTTAAGACGCAAGCTGAGAACTCT	12584
Qy	591	ArgGluGlyIleThrProTrpAlaSerPheLysLysMetValThrProLysLysArgVal	610
Db	12585	AGGGAAGAACTTAGCGCAGAGTCGGTGCCGCCGAAAAATCACCTTTGCCCTCCAAAGAAGCT	12644
Qy	611	ArgArgProSer-----GluSerAspLysGluLuGlu-----	621
Db	12645	TCGAGACCAACTTCGTTGCAGAAAGTGTAAAGNGTAGAGCTGATAGTCCAAAGAGGAA	12704
Qy	622	-----LeuGluLysVallYsSerAlaThrLeu-----SerSerThr	633
Db	12705	TCTAGCGCAGAGTCGGGGCCGNAATACTCTTTGGCTCCATGGAAGCTTCCAGACCA	12764
Qy	634	AspSerThrValSerGluMetGlnAspGluValLysThrValGlyGluGlu-----	650
Db	12765	ACTTCGTTGCAGAAAGTGTTAAGACCAAACCTGAGAAAGTCTAAGGAAGAATCTAGCGCA	12824
Qy	650	-----	650
Db	12825	GAGTCGGTGACCGAAAAATCTCTTTGCCCTCCAAGGAAGCTTCCAGACCAACTTCGTT	12884
Qy	651	-----GlnLysProGluGluProLysArgArgValAspThr	662
Db	12885	GCAGAAAGTGTTAAGGATGAAGCTGAGAAGTCTAAGGAAGAATCTAGCGCA-----GAG	12938
Qy	663	SerValSerTrpGluAlaLeuIleCysValGlySerSerLysLysArgAlaArgLysAla	682
Db	12939	TCGGTGGCCGAAAAATCTCTTTGCCCTCCCAAGGAATCTCCAGACCAAGCTTCGCTGCA	12998
Qy	683	SerSerSerAspAspGlu-----	688
Db	12999	GAAAGTATTAGGACCAAGCTGAGGGGACTAAGCAGGAATCTAGCGCAGAGTCGATGCC	13058
Qy	689	-----GlyGlyProArgThrLeuGlyGlyAspSerHisArgAlaGluGluAlaSerLys	706
Db	13059	GAAACGGTGAAGCGCAGAGATCAAGGGTGTAT-----CAATCTCTTTAGCGTCCAAA	13112
Qy	707	AspLysGluAlaGlyThrAspAlaValProAlaSerThrGlnGluGlnAspGlnAlaGln	726
Db	13113	GAA----ACGTCAAGACCCGATCTCTTGTGAAAGACGTTAAGGACGAACTGAAAAAGCCA	13169
Qy	727	GlySerSerProGluProAlaGlySerProSerGluGlyGlyGlyValSerThrTrp	746
Db	13170	GNAGGTCAGCATCGACAGTCGAGGTGGCATCCGACAGAGTCCGTGGCAGTA---13226	
Qy	747	GluSerPheLysArgLeuValThrProArgLysLysSerLysLysLeuGluGluLys	766
Db	13227	--ACTGCTAAGGACGAGAAGTCACCGCTTCATTTCGGTCTCTGAATCAGTTGCCGACAG	13283
Qy	767	AlaGluAspSerSer-----ValGluGlnLeuSerThr	777
Db	13284	TCGCAGATGCTTCGAAGGAAGCATCTCGTTCTTTATCGGTTGGCGAAACCGCCTCATCA	13343
Qy	778	GluIleGluProSer--ArgGluGluSerTrpValSerIleLysLysPheIleProGly	796
Db	13344	CCCATTGAAGAGGTCCACGATCCATTCTGACCTCAGTCTGCCCCCTTAAACCTTACAGGT	13403
Qy	797	ArgArgLysLysArg-----AlaAspGlyLysGln	806

Db	13404	GAAGCTAAGGCGAAGCTTCCAACCTCTTTCGAGTCGCCAATTGATGTGGCAGAAGGAGACTTT	13463
Qy	807	GlulGlnAlaThrValGluAsp-----SerGlyProValGluIle	819
Db	13464	TTAGAGTAAGCGCGAATCTCTCCCGCGGCCAGCTGTACTTTCAAAACCCAGCGGAGTTT	13523
Qy	820	AsnGluAspAspProAsnValProAlaValValProLeuSerGluTyrAsnAlaVal---	838
Db	13524	TCGCGCCGACTACTGGGCATACAGCTCCAGCCCGTGGACGAGGCTTCTCCCGTTCTG	13583
Qy	839	----GluArgGluLysMetGlu-----AlaGlnGlyAsnThr	849
Db	13584	GAGGAATTCGAAGTAGTTGAACAGCATACCACTTCTCGAGTCGGGGCTACGGCGCAACA	13643
Qy	850	GlulLeuProGlnLeuLeuGly-----AlaValTyrValSerCluclulLeu	864
Db	13644	CGAGAAACGGACTTACTTGACCTGACTGAACCAAAATCCGACACTGTAAACCAAGCAATCG	13703
Qy	865	SerLysThrLeuValHisThrValSer-----ValAlaValIleAsp	878
Db	13704	GAGACAACGTTATTTGAACCCCTTACCTCGAAGTTGAAGTTGAAGTGGAGTACTGGAG	13763
Qy	879	GlyThrArgAlaValThrSerValGluGluArg-----SerProSerTrpIleSer	895
Db	13764	AGC-----TCGGTGAACAAGTTGAAGAGAAAGTCCAGACCTCAGTAAAGCAAGCAGAA	13817
Qy	896	AlaSerValThrGluProLeuGluHisThrAlaGlyGluAlaMetProProValGluGlu	915
Db	13818	ACCACCTGTGACCGATTCCTTGGAAACAACCAACCAAGAGAGAGCAGTGCAGTGTGACAG	13877
Qy	916	Val-----ThrGluLysAspIleAlaGluGluThrProValLeuThr	930
Db	13878	ATCAAAATCAGTGTAGACACCACTTTGAGAGAGTAGCGAAATCGTTGCCGATGTGGCC	13937
Qy	931	GlnThrLeuProGluGlyLysAspAlaHisAsp-----AspMetValThr	945
Db	13938	AAAGTACTTAAGTCTGCAAGAGACATTACAGACATCATACCCGACTTTTGACGAACGGCAA	13997
Qy	946	SerGluValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeuArg	965
Db	13998	CTGGAGGAGAAACTGAAGTCAACTCCGATACAGAGGAGGATCCGACAAAGTAGTACCCGG	14057
Qy	966	ThrGluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSerAlaVal	985
Db	14058	GACGAGAAATCCCTCGAAATCAGCGTCAAGTAGAA-----ATC	14096
Qy	986	SerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGluValGlu	1005
Db	14097	GAATCTGAGAGTCTATCGCCCGAT-----CAGAAAGTCCGGCCCATTTCCATCGAGGAG	14150
Qy	1006	SerGlyValLeuAspThrGluGluGluArgGlnThrGlnAlaIleLeu-----	1022
Db	14151	AAGGACAAATTCAGCAATTCAGAAAAACACAGTGTAGCGAGGGAATATTGACCAAGCAGT	14210
Qy	1023	-----GlnAlaValAla-----AspLysValLysGluGluSerGlnValProAla	1037
Db	14211	CGCGCGAGTCHGTTTGCCAGCCAGCGGAGTGGTGCCCTTCCCTTCCAGTCCAGCAGCC	14270
Qy	1038	ThrGlnThrValGlnArgThrGlySerLysAlaLeuGluLysValGluGluValGluGlu	1057
Db	14271	AGCCATGAACATAAGGAGGTGGAGTATCCGAGAGCCATAAAGCTGAGAGTCTTCTCGC	14330
Qy	1058	AspSerGluValLeuAlaSerGluLysGluLysAspValMetProLysGlyProValGln	1077
Db	14331	CCCGAGTCCGTTGCCAGCCAGGTCCAGCGAAAAAGGACATGAAGACATCCGCGCCAGCATCC	14390
Qy	1078	GluAlaGlyAlaGluHisLeuAlaGlnGlySerGluThrGlyGlnAlaThrProGluSer	1097
Db	14391	AGCACCGCCCAATTTAGCACCAAGGAAGCGCATGAG-----GAGACCACTGAGTCG	14441
Qy	1098	LeuGluValProGluValThrAlaAspValAspHisValAlaThrCysGlnValIleLys	1117

Db 14442 CTA-----CTTCACTGCTTACTACCAACAGACGACGTGGAG 14477
 Qy 1118 LeuGlnLeuMeuGlnValAlaProGluSerSerGluThrLeuThrAspSer 1137
 Db 14478 ACGAACAAGATGAGAGAAATCGACGTTGATGTCCTCCACTCACTCACTAAGTCC 14537
 Qy 1138 -----GluThrAsnGlySerThrProLeu 1145
 Db 14538 ACAGTCTCTCCAGTCAGTCAACTGTTCACTTCGAGAAAGCTCCACGACGAAATCGTTG 14597
 Qy 1146 AlaAspSerAspThrAlaAspGlyThrGlnAspGluThrIleAspSerGlnAspSer 1165
 Db 14598 AGTAGCAAGTCTCAAGTGGAGACACAGTCCCGACGAGTCCGCTTCACGCTTGTTGGGG 14657
 Qy 1166 LysAlaThrAlaAlaValAlaArgGlnSerGlnValThrGluGlnGluAlaAlaThrAla--- 1184
 Db 14658 GAGAAAGAGAGATCGCTACCAACTAGTTTGAAGAGACACTAGCCGCTCGATCTC 14717
 Qy 1185 -----GlnLysGluGluProSerThrLeuProAsnAsnValPro 1197
 Db 14718 CAGCTGGAGAGAGCTGTTGTGCGAGGAAATGCTCTCAAGTCAATCGTTAGCGAG 14777
 Qy 1198 AlaGlnGluGluHisGlyGluGluProGlyArgAspValLeuGluProThrGlnGlnGlu 1217
 Db 14778 ATTCACACACAGCATCGCTCAGAAATCAATTAAGAAATCAAGATCTAGAGAAACCAA 14837
 Qy 1218 LeuThrAla-----AlaAlaValProValLeuAlaLysThrGluValGlyGlnGlu 1234
 Db 14838 GTAAACGACGACGATTACCACTACACACAGAGCCACCAAAAGCATAGCTTAAAGAA 14897
 Qy 1235 GlyGluValAspThrLeuAspGlyGlyLysValLysGluGlnGluValPheValHis 1254
 Db 14898 ACTGTGCGAGATTGTCGACCGACGAGAAATGTTGTCGCGCAAGAGCGTTC----- 14951
 Qy 1255 SerGlyProAsnSerGlnLysAlaAlaAspValThrTyrAspSerGluValMetGlyVal 1274
 Db 14952 ---AGACAGTGAAGCCACTAAGTGGCCGAC----- 14978
 Qy 1275 AlaGlyCysGlnGluLys---GluSerThrGluValGlnSerLeuSerLeuGlnGly 1293
 Db 14979 ---GACTGCTTGAAGAAAGACACCGCAGCGCGCTCCAGACATCCGCTCCAGCGA 15035
 Qy 1294 GluMeuGluThrAspValGluLysGlyLysArgGluThrLysProGluGlnValSerGlu 1313
 Db 15036 GCGCTATTGCTAGGACGCGAGCAATCACGGAGAGAGTCTGCTGCTGACGAGCAAGTGA 15095
 Qy 1314 -----GluGlyGlnGlnGluThrAlaAlaProGluHisGlu-GlyThrTyrGly 1329
 Db 15096 AGCCGCTTGAACCCACAGTATCTCTGAAGATGAGAAACCGCTATGATACGTGCGAGCGG 15155
 Qy 1329 LysProValLeuThrLeuAspMetPro----- 1338
 Db 15156 T---CCAGTGTTAAGAGAGCGCTTCCAAAGTATTGCTACATCATGATGACAGCATC 15212
 Qy 1339 -----SerSerGluArg-GlyLysAlaAla 1346
 Db 15213 TACAAACCTTCCAGAGATATGAAACGATTTCAAAGCTCTCGAGAGAGACAGGACAC 15272
 Qy 1346 euGlySerLeuGlyLysSerProSerLeuProAspGlnAspLysAlaGlyCysIleGluVal 1366
 Db 15273 GTGGAGAGAGTGGCTCAGAGAGTCACTTCAACAGAAAGCAGC---AACTCTTTGGCAA 15329
 Qy 1366 AlaGln-ValGlnSerLeuAspThrThrValThrGlnThr-----AlaGluAlaVal 1382
 Db 15330 TCCAGGGAACAGAGACAGCAACCAACAGTAGACACCAAGACCGGTGCTCAAGGTTG 15389
 Qy 1383 GluLysValIleGluThrValValIleSerGluThrGlyLysSer----- 1397
 Db 15390 GAGTCCATTACCTTCAACAGATGATCAAGAAACTAGGACGAGTCAAGGCGACCTTGCC 15449
 Qy 1398 -----ProGluCysValGlyAlaHisLeuLeuProAlaGlyLysSer---Ser 1412
 Db 15450 GATGCAAGACACCGGCACT-----GCGCCAGTAAAGTCCCGGCGTCAAGCAATGAGC 15503

Qy 1413 AlaThrGlyGlyHisThrProLeuGlnHisAla---GluAspThrValProLeuGlyPro 1431
 Db 15504 TCGACAGATACACCGGCTCTGTGATTTAGAGCTGAGCTGAGACCGCTGGTGA 15563
 Qy 1432 GluSerGlnAlaGluSerIleProIleLeuValThrProAlaProGluSerThrLeuHis 1451
 Db 15564 AAGTGTAGTCAGAGTCCCGCCAGCATCTGATCTTCAATCGGGCCGATGTCGCC- 15617
 Qy 1452 ProAspLeuGlnGlyLysLysSer-----AlaSerGlnArgGlu----- 1464
 Db 15618 AAGACATACAGCGGCAAACTTAGCCCGCTGCTCAGCTCGAGAGACAGCAATTC 15677
 Qy 1465 -----ArgSerGluGlnLysProAspAlaGlyProAspAlaAspGly 1480
 Db 15678 ACTCGCTAGACGCGGAGTTCGATACGACACCCCGAATCTCGCCA----- 15725
 Qy 1481 LysGluSerThrAlaIleGluLysValLeuLysAlaGluProGluIleLeuGluLeuGlu 1500
 Db 15726 AAGCCACATCACCTTTCGCGGCTCAGACAGACAGCTTAAGTCTTGAAATGCA 15785
 Qy 1501 SerLysSerAsnLysIleValLeuAsnValIleGlnThrAlaValAspGlnPheAlaArg 1520
 Db 15786 GTATGTTAAGCAGTTTGTGCTGATTTATTTGCTATTATATCTATGTTTAAAC 15845
 Qy 1521 ---ThrGluThrAlaProGluThrHisAlaTyrAspSerGlnThrGluValPro----- 1537
 Db 15846 CCATGTACACCACTCGCAGAGCAGATGTGGCAGAGACGCTGCTCAGCCGAGCTG 15905
 Qy 1538 AlaCysArgLeu-AspSerArgGluProAsnArgCysThrProLysMetLysAspAlaLys 1557
 Db 15906 AGTGTAGGAGTACATTCGAGCTGCACGAACTCGCTGACTAAGTGC----- 15954
 Qy 1557 SwetLysHisProValProGlnProArgLysAspLeuGln 1570
 Db 15955 -----CACACAGCCCTCAG---CGGAGACACCGACAA 15984
 RESULT 12
 ABL19989
 ID ABL19989 standard; DNA; 20448 BP.
 AC ABL19989;
 DT 26-MAR-2002 (first entry)
 DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 11440.
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ds.
 XX Drosophila melanogaster.
 OS
 PN WO200171042-A2.
 PD 27-SEP-2001.
 PF 23-MAR-2001; 2001WO-US09231.
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 PA *PEPE) PE CORP NY.
 PI Venter JC, Adams M, Li PWD, Myers EW;
 DR WPI; 2001-656860/75.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 PS Claim 1; SEQ ID NO 11440; 21bp + Sequence Listing; English.
 XX

CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABBS7737-ABBS7072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 20448 BP; 7528 A; 4083 C; 4878 G; 3959 T; 0 other;

Alignment Scores:

Pred. No.: 3,9e-19 Length: 20448
 Score: 557.50 Matches: 396
 Percent Similarity: 36.48% Conservative: 300
 Best Local Similarity: 20.75% Mismatches: 725
 Query Match: 6.91% Indels: 487
 DB: 23 Gaps: 79

US-09-902-432-4 (1-1596) x ABL19989 (1-20448)

QY 7 ThrGluGlnArgSerProGluGlnProAlaGly----- 17
 DB 8965 ACTACAAACGAATGTTGCGACGCTGCTGGCGAAGGTGAATAGAGATTATGAAGTT 9024
 QY 18 -----SerAspThrPro-----SerGluLeuValLeuSerGlyHisGlyProAlaAlaGlu 34
 DB 9025 GTTAGGGATGATCAGCTGAGGCGAGAAATACTATTGTGAGTATGAGCCA----- 9075
 QY 35 AlaSerGlyAlaAlaGlyAspProAlaAspAlaAspProAlaThrLysLeuProGlnLys 54
 DB 9076 -----GAACTGTGAACCAAGACGMAAAACCCAAAGAACTTAAGAG 9117
 QY 55 AsnGlyGlnLeuSerSerValAsnGlyValAlaGluGlnGlyAspValHis-----Val 72
 DB 9118 AAACCCAGG-----AAGGTTAAAGAAAGATGATATACATGACTACATA 9159
 QY 73 Gln-----GluGluAsnGlnGluGlyGln 80
 DB 9160 CAAAGTTAATTGAGTTGGAAACCAACCAAGACGAGCTTGAAAGATATGAAATTTGAA 9219
 QY 81 GluGluGluValValAspGlu-----AspValGlyGlnArgGlu 93
 DB 9220 TTCCGAACCAATAGTCAAGGACAGCCATTTGGATAGCCGATTCACGTTTGGACGAATCG 9279
 QY 94 SerGluAspValArgGluLysAspArg----- 102
 DB 9280 CCAAAAGAGTGCAGAAAAGATATAAAGTCAAGGTCTCAAAAGTTCCGAATGAGGAA 9339
 QY 103 -----ValGluGluMetAlaAlaAsnSerThrAlaValGluAspIleThrLysAspGly 120
 DB 9340 ACTCTGTTCCAGAACAAATATGCAAGGTCAATGTTGTGCAAGAAAGCTCTCTGAGCAA 9399
 QY 121 GlnGluGluThrSerGluIleGluGlnIleProAlaSerGluAsnValGluGlu 140
 DB 9400 CCGGAATACCAAGTGCAGAAATTCAGAAAGTTAAGCCCGTA---GAGGTAGATGTTAAAGAG 9456
 QY 141 MetValGlnProAlaGluSerGlnAlaAsnAspValGlyPheLysLysValPheLysPhe 160
 DB 9457 GTTATTAAGTGAAGATGCGAAGCGGTTCAAGAAAAGAAAGACTACTAAGCGAGTTTAAAGAG 9516
 QY 161 ValGly-----PheLysPheThrValLysLysAspLysAsnGluLys 174
 DB 9517 ATTGGCCCGGAGAACAAACCACTTTAAGATTACAAATG-----ATTGAAGCCGAGGAC 9570
 QY 175 SerAspThrValGlnLeuLeuThrValLysLysAspGluGlyGlyAlaGluAlaSer 194
 DB 9571 AACGATTCTGTT-----ACAGTTATTGTTGACGAGAACCGGAAATAGCTTCTCCA 9621
 QY 195 ValGlyAlaGlyAspHisGlnGluProSerValGluThrAlaValGlyGluSerAlaSer 214

DB 9622 CAGTCTATCGAAGAACATCCGAAACAGTCTTAAGGAAAGAGTCTCAAGGCCCAAGAAA 9681
 QY 215 -----LysGluSerGluLeu----- 219
 DB 9682 ACTGTTCAGAAAAGTTTAAAGAGCAGCTTAAAGGATTAAGGATTAAGGATTAAGGAA 9741
 QY 220 -----LysGlnSerThrGluLysGlnGluGlyThrLeuLysGlnGluGlnSer 236
 DB 9742 GAAATTCCTTAAGGTTGACCTTGAGAAATATGAA-----AAGGTAGAAATGCCAGAA 9792
 QY 237 ThrGluIleProLeuGlnAlaGluSerAspGlnAlaAlaGluGluGluAlaLysAspGlu 256
 DB 9793 AAACGATTAAACTACTGTT---TCTGATTCCATTCGGAAGAACCAAACTGACAAA 9849
 QY 257 GlyGlu-----GluLysGlnGluLysGluProThrLysSerProGlu 270
 DB 9850 TCCCAACCAATATCAGTACTACCCGACACACGAAACCCCAAGAAACGAAGACTCCCAAG 9909
 QY 271 SerProSerSerProValAsnSerGluThrThrSerSerPheLysLysPheThrHis 290
 DB 9910 ACTCCAAAGACAGACACAGACACCAAGTTCCGACGAAACCCGACAGACCACTGTA 9969
 QY 291 GlyTrpAlaGlyTrpArgLysLysThrSerPheLysLysSerLysGluAspLeuGlu 310
 DB 9970 GACACTACTGACATACAGAGCTTACCAACCCCAACTGCACAGCCCGGACACACAGCC 10029
 QY 311 ThrAlaGluLysArgLysGluGlnGluAlaGluLysValAspGluGlu----- 327
 DB 10030 ACTGCACAATTTACCAACGCGCACCAAGAGAGAATCTTACTCAAGTACACTAAAGAT 10089
 QY 328 -----LysGluLysThrGluProAlaSerGluGluGlnGluPro 340
 DB 10090 ACAATTCAGAAAACAGTTAAACATAAGAAAACAAACCA-----GACACACAAAAAGC 10143
 QY 341 AlaGluAspThrAspGlnAlaArgLeuSerAlaAspTyr----- 353
 DB 10144 GTTGAACCAAGTGAATACCAAGGTTTCAATGAAGACTATCAAAATTAGCATCATCCAGAA 10203
 QY 354 -----GluLysValGluLeuProLeuGluAspGlnValGlyAspLeuGlu 368
 DB 10204 GAGCTTGTTCGAGGAGAGAGAGCCGCGGAAAGATTTTGGAGGTTAGAGTTATCGAT---GAA 10260
 QY 369 AlaSerSerGluLysCysAlaProLeuAlaThrGluValPheAspGlu----- 385
 DB 10261 GTTCGCGAGGTGGAAGAATCACAGCCCAATTGTGGAAGAGGTAGAAGTACAGAACACACAG 10320
 QY 385 ----- 385
 DB 10321 CCAGCTACAGAAAGAACTGTTGAGGACGTCCTAAGCTTAAAGTCTAAAAAGAAAGAGTG 10380
 QY 386 -----LysMetGluAlaHisGlnGluValValAlaGluValHisValSerThrVal 402
 DB 10381 GTTAAAGAAAGACCGATGACCATGACGACTCATCAAAAGATGTTGGAGCAGGAAATA 10440
 QY 403 GluLysThrGluGluGlnGlyGlyGlyGluAlaGluGlyValValValValGlu 422
 DB 10441 GAGAAGACGAACTCGAAAAATACGAGAAAATTGAATTTGATGTTCCCAAGAAACTCAAA 10500
 QY 423 GlyThrGlyLysSerLeuProGluLysLeuAlaGluProGlnGluValPro----- 440
 DB 10501 CCAGAAATTCGCTCTCTCGAACCAATAAAATAGAGCGCAAGAACAGAACCAAAAG 10560
 QY 441 -----Gln 441
 DB 10561 GTGACTATTCTCGATGCCACTGATGTTCTTAAACCGTGAAACTAAACCGTAGTAACGT 10620
 QY 442 GluAlaGluProAlaGluGluLeu-----MetLysSerArg 453
 DB 10621 AAGGAGAAACCTGCGGAGAACTGACCGTTCACTTCCAAAGTTCCAGACTTAAGCCCGT 10680
 QY 454 GluMetCysValSerGlyGlyAspHisThrGlnLeuThrAspLeuSerProGluGluLys 473

QY 1096 GluSerLeuGluValProGluValThrAlaAspValAspHisValAlaThrCysGlnVal 1115
Db 12661 GTGTTAGTGAAGCTCTGAGAGTAAGATTGTAGAGGAAGTATTTCAGAGAAG 12720
QY 1116 IleLys-----LeuGlnGlnLeuMetGluGlnAlaValAlaProGlu-----SerSer 1131
Db 12721 CCTAAGAGTTCAATTCGGTTTCTGAATCAGAACCCAGCCGAGGAGCCAGTGTC 12780
QY 1132 GluThrLeuThrAspSerGluThrAsnGlySerThrProLeuAlaAsp----- 1147
Db 12781 GAGCAGTTACAGTAAATAAAGCAAGCCCTCTGTTACTTTCGCCGACGAACCTGCCACT 12840
QY 1148 -----SerAspThrAlaAspGlyThrGlnGlnAspGluThrIleAsp 1161
Db 12841 GAAATCGTTATCAGGAAGAACGCGCTGAAGTTGCTACTGAAGACGCTCATATCAAG 12900
QY 1162 SerGlnAspSerLysAlaThrAlaAlaValArgGlnSerGlnValThrGluGluAla 1181
Db 12901 ACCAAGAGCCAAAG-----AAGAGGTAAACGGATGTGGAAGCT 12939
QY 1182 AlaThrAlaGlnLysGluGluProSerThrLeuProAsnAsnValProAlaGlnGluGlu 1201
Db 12940 GAGGAGCTAAAGATTAAAGTCACTGAGGAAGTCCACAGGAATTCCTCATATTGGAGGAG 12999
QY 1202 HisGlyGluGluProGlyArgAspValLeuGluProThrGlnGlnGluThrAlaAla 1221
Db 13000 GTCTCGGAGGAA-----GAGGTGCTACTGAAACCAAGAG-----ACA 13038
QY 1222 AlaValProValLeuAlaLysThr--GluValGlyGlnGluGlyGluValAspTrpLeu 1240
Db 13039 GCGCTCTGTTGAGAGAGACATACAAATTTGGTATCAAGGAACCGGAGCCGGAAG 13098
QY 1241 AspGlyLysValLysGluGlnGluValPheValHisSerGlyProAsnSerGln 1260
Db 13099 CTGCGCAAGCAATTGTCGAGGAGGAGCGCGTGTACAGAGCCCAATTCAGGAAGCA 13158
QY 1261 LysAlaAlaAspValThrTyArgSerGluValMetGlyValAlaGlyCysGlnGluLys 1280
Db 13159 CCGAAACCGGAAGTCTTTGAGGAACACACAGTTAGATCATT----- 13200
QY 1281 GluSerThrGluValGlnSerLeuGluGlyGluMetGluThrAspValGlu 1300
Db 13201 -----GAGGAACACCTCGGGAATTTGTCAGGAAGTTCATTGAGGAGGAGGTGAAG 13251
QY 1301 LysGluLysArgGluThrLysProGluGlnValSerGluGlyGluGlnGluThrAla 1320
Db 13252 GTAATTCGAGAAAGAGCTTAAGCAGAAATTAAGAGGAACCCGAGCGCGAGTTACT 13311
QY 1321 AlaProGluHisGluGlyThrTyArgProValLeuThrLeuAspMetProSerSer 1340
Db 13312 GTT-----TCTACTCCAAAGCCTGTC----- 13332
QY 1341 GluArgGlyLysAlaLeuGlySerLeuGlyGlySerProSerLeuProAspGlnAspLys 1360
Db 13333 GAAGAAGTGAAGCTACGTCAGATTGCTGTTATTCAGAACCAACCCGACTGAGGAGGAA 13392
QY 1361 AlacGlyCysIleGluValGlnValGlnSerLeuAspThrValThrGlnThrAlaGlu 1380
Db 13393 GCCGCTGATCTCAGATAACATCATTTGAGAAGAA-----ACACCACCAAGAA 13443
QY 1381 AlaValGluLysValIleGluThrValValIleSerGlu----- 1393
Db 13444 CTAGTTCCAGGAATTTGAAGAAATTTGAATTCGTTGAGGAACCAAGGCCCTGAAGAACAA 13503
QY 1394 -----ThrGlyGluSerProGluCysValGlyAlaHisLeuLeu 1406
Db 13504 CCAACGGATTTCACCTTCGCCACAAAGGATTCGAAAGAAAGCCGACTGTGGAAGAGCTG 13563
QY 1407 ProAlaGluLysSerSerAlaThrGlyGlyHisTrpThrLeuGlnHisAlaGluAspThr 1426
Db 13564 CCCAGGACACAGTGC-----ACGATTCAAAAAAAGAAAGAAAG 13602
QY 1427 ValProLeuGly-----ProGluSerGlnAlaGluSerIleProIleIleValThr 1443

Db 13603 GCACCAAGTACCAAGAGTAGTTGAAGAACCCGAGCTGAGTTTGTGTTAAGCCAAAGACG 13662
QY 1444 ProAlaProGluSerThrLeuHisProAspLeuGlnGlyGluIleSerAlaSerGlnArg 1463
Db 13663 CTTGTCCAAGAGTTACT-----GAAGAAGCTAAGATAACGAAGTCTAAGAAA 13710
QY 1464 GluArgSerGluGluGluAspLysProAspAlaGlyProAspAlaAspGlyLysGluSer 1483
Db 13711 CTTGTAAGGAAGAGAG-----GCCGCGGAGAGCTTAAA 13746
QY 1484 ThrAlaIleGluLysValLeuLysAlaGluProGluIleLeuGluLeu----- 1499
Db 13747 GTAACGATCACTGAAGAAATTCCTCACTGAACCAAGATTCAAGAAATTTTCGAGGAGATC 13806
QY 1500 -----GluSerLysSerAsnLysIleValLeuAsnValIleGlnThrAlaVal 1515
Db 13807 GAAGAGATCGAGGAAGAAAACCCGAGAGTATGTGATCGAAGTTAAGGAAGCCCAACCA 13866
QY 1516 AspGlnPheAlaArgThrGluThr-----AlaProGluThr 1527
Db 13867 GAAGCTGTAGAGATAAGGAAGTTAGCTTACCAGAAAAGAAACCTAAGGCTCCAATAGTT 13926
QY 1528 HisAlaTyArgSerGln---ThrGlnValProAlaCysArgLeuAspSerArgGluPro 1546
Db 13927 GAAGACCCAGAGCTGAATAAATTTCAAGCCAAAGGTAAAGTCGGAAGAAAGTTCAGGAA 13986
QY 1547 AsnArgCysTrpThrLysMetLysAspAlaLysMet----- 1558
Db 13987 GAAGCAAGATAGTAGAAGAAAAGCCAAAGAAAATCGACGAAGTTGCAGTAGCCGAGCAG 14046
QY 1559 -----LysHisProValProGlnProArgGluAspLeuGlnValLeu 1572
Db 14047 CTTACGCTCAAGGTTCAAGAAGAGGTTGTCCTCAGAGCCGATTCTTGAAGAAGAAAGTGATA 14106
QY 1573 ThrValLeuGluAlaTrpAlaGlnProArgLysCysLeuProArg----- 1587
Db 14107 GAAGAGTTCCAAATTTAAAAAGAAACCAAAAGAACCCGAGCAGAGGACATTTGTCGATGCT 14166
QY 1588 -----LeuGlnLeuLysAlaPro 1593
Db 14167 GCGATTGTGAATCAAGAAACCA 14190
RESULT 13
ABL19988/c
ID ABL19988 standard; DNA; 24971 BP.
XX ABL19988;
AC
XX 26-MAR-2002 (first entry)
DT
XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 11437.
DE
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
KW
XX Drosophila melanogaster.
OS
XX WO200171042-A2.
PN
XX 27-SEP-2001.
PD
XX 23-MAR-2001; 2001WO-US09231.
PF
XX 23-MAR-2000; 2000US-191637P.
PR
XX 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers BW;
XX
DR WPI; 2001-656860/75.
XX

Qy	544	SerAlaAspGluGlnLysGlyGluSerSerAlaSerSerProGluGluProGluGluThr	563
Db	10534	-----GAAAGAAAGACGTCGAGAAACCAAAATCTTACGAGTTCAAAATTTACAGACG	10481
Qy	564	ThrCysLeuGluLysGlyProLeuGluAlaProGlnAspGlyGluAlaGluGluGlyThr	583
Db	10480	CMATCTATTGAAGAAACCAATAGAGTGCTGTAAGAAAGCTCTCGAGGAA-----	10430
Qy	584	ThrSerAspGlyGluLysLysArgGluGlyIleThrProTrpAlaSerPheLysLysMet	603
Db	10430	-----	10430
Qy	604	ValThrProLysLysArgValArgProSerGluSerAspLysGluGluGluLeuGlu	623
Db	10429	---ACACCAAAAGTTGTAGAGAAAAGGTGCCGAGAAATTCGATTCCTACGAATTC---	10376
Qy	624	LysValLysSerAlaThrLeuSerSerThrAspSer-----ThrValSerGlu	639
Db	10375	-----ACTCTAAAGAACTGATGAAGAAAAGGTAAATCACAGTTTGACGAC	10333
Qy	640	MetGlnAspGluValLysThrValGlyGlu-----GluGlnLysProGluGluProLys	657
Db	10330	CAGCCCGAGGAAGAGCCCTGTAGAGTTGTATTCAAGAAAACCCCAAGAACCTTGAA	10271
Qy	658	ArgArgValAspThrSerValSerTrpGluAlaLeuIleCysValGlySerSerLysLys	677
Db	10270	GCAGTTGAGCGCCAGTTTGTATGATCGAACCCCAAGATTGTGGAAGAAACGAGCGTGAA	10211
Qy	678	ArgAlaArgLysAlaSerSerSerAspAspGluGlyGlyProArgThrLeuGlyGlyAsp	697
Db	10210	ACTGCAATCAAA-----	10199
Qy	698	SerHisArgAlaGluGluAlaSerLysAspLysGluAlaGlyThrAspAlaValProAla	717
Db	10198	CAGAAAAAGACGAAAAAACCCGAGAGAGACGAGGAGAGCTCAATTAGCAATCAAGTT	10139
Qy	718	SerThrGlnGluGlnAspGlnAlaGlnGlySerSerSerProGluProAlaGlySerPro	737
Db	10138	GTCCAAAGTGGAAGCCCGGTTGCCGAAGAGTGTTCAGTGAAGCTCCTGAGAGTAAGATT	10079
Qy	738	SerGluGlyGluGlyValSerThrTrpGluSerPheLysArgLeuValThrProArgLys	757
Db	10078	GTAGAGGAGGAAGTTATTGCA-----	10058
Qy	758	LysSerLysSerLysLeuGluGluLysAlaGluAspSerSerValGluGlnLeuSerThr	777
Db	10057	-----GAGAAAGCTTAAGAGTTTCACAAATTCG-----GTTTCT	10022
Qy	778	GluLeuGluProSerArgGluGluSerTrpValSerIleLysLysPheIleProGlyArg	797
Db	10021	GAATCAGAAACCAAGCCCGAGGAGCA-----AGTGTCGAGCAGTTTCACAGTTTAAAAAA	9968
Qy	798	ArgLysLysArgAlaAspGlyLysGlnGluGlnAlaThr-----ValGluAspSer	814
Db	9967	CGCAAGCCTCTGTGTACTTTCGCCGACCAACCTGCACATGAAATCGTTATTCAAGGAAAGC	9908
Qy	815	GlyProValGluIleAsnGluAspAspProAsnValProAlaValValProLeuSerGlu	834
Db	9907	AAGCCCGCTGAAGTTGTCTACTGAAGACCGCTCATATCAAGACCAAGACCAAGAAAGAG	9848
Qy	835	TyrAsnAlaValGluArgGluLysMetGluAlaGlnGlnGlyAsnThrGluLeuProGlnLeu	854
Db	9847	GTAACGGATGTGGAAGCTGAGGAGCTAAAGATTAAAGTCACTGAGGAAGTTCCACAG---	9791
Qy	855	LeuGlyAlaValTyrValSerGluGluLeuSerLysThrLeuValHisThrValSerVal	874
Db	9790	-----GAAATTTCCATATTGAGGAGGCTCTCGAGGAA-----	9758
Qy	875	AlaValIleAspGlyThrArgAlaValThrSerValGluGluArgSerProSerTrpIle	894
Db	9757	GAGTGATCACTGAAMACCAGAAAGACACGGCCTGTGTGTTGAAGAGAGACATACAAATTT	9698
Qy	895	SerAlaSerValThrGluProLeuGluHisThrAlaGlyGluAlaMetProProValGlu	914

Db	9697	GGTATCAAGGAAACGGAGCCGGAAG--			-----CCTGCCGA	9662
Qy	915	GluValThrGluLysAspIleIleAlaGluGluThrProValLeuThrGlnThrLeu---			-----	933
Db	9661	GCAATTGTGAG--			-----	9617
Qy	934	-----ProGluGlyLysAspAlaHis-			-----	940
Db	9616	GAAGCACCGAAACCGGAAGCTTTTGGAGAACCAAGGTTAGAGTCATTGAGGAACACCT			-----	9557
Qy	941	-----AspAspMetValThrSerGluValAspPheThrSerGluAlaValThr			-----	956
Db	9556	CGGAAATTTGGTCGAGGAAGTCATTGAGGAGGAGGTCAAGGTAATTCGCGAAGAAAGCCT			-----	9497
Qy	957	AlaThrGluThrSerGluAlaLeuArgThrGluGluValThrGluAlaSerGlyAlaGlu			-----	976
Db	9496	AAGCAGAAATTAAGAGGAACCCGAGCGGAGTACTGTTTCTCTCAAAGCGCTGC			-----	9437
Qy	977	GluThrThrAspMetValSerAlaValSerGlnLeuThrAspSerProAspThrThrGlu			-----	996
Db	9436	GAAGAGTGGAGCTAGCTCCAGTATTCGTTATTCAGAACACCG--ACTCAGGAG			-----	9380
Qy	997	GluAlaThrProValGlnGluValGluSerGlyValLeuAspThrGluGluGluGluArg			-----	1016
Db	9379	GAAGCGCTGATCTCAAG--			-----ATAACAATCATTTGAAGAAACA	9338
Qy	1017	GlnThrGlnAlaIleLeuGlnAlaValAlaAsp--LysValLysGluGluSerGlnVal			-----	1035
Db	9337	CCACCACAAGAACTAGTTCAGGAATTTGAAGAAATGTAATTCGTTCGGAACCAAGGCC			-----	9278
Qy	1036	ProAlaThrGlnThrValGlnArgThr--GlySerLysAlaLeuGluLys-			-----	1051
Db	9277	CCTGAGAACCAACCGGATTTCACTTCGCCACAAAGGATTCGAAAGAGCCGACT			-----	9218
Qy	1052	ValGluGluValGluGluAspSerGluValLeuAlaSerGluLysGluLysAsp----			-----	1069
Db	9217	GTGGAAGAGTGCCTGAGGAACAGGTTCACGATTCAAAAGAAAAAGAAAGACCAAGTGA			-----	9158
Qy	1070	-----ValMetProLysGlyProValGln			-----	1077
Db	9157	CCAGGAGTGTGAAGAACCCGAGCTGAGTTTGGTTTAAAGCCAAAGCCCTGTCGA			-----	9098
Qy	1078	-----GluAla			-----	1079
Db	9097	GAAGTTACTGAAGAGCTAAGATAACGAAGCTCTAAGAAACCTGTAAAGCAAGAGGCC			-----	9038
Qy	1080	GlyAlaGluHisLeuAlaGlnGlySerGluThrGlyGlnAlaThrProGluSerLeuGlu			-----	1099
Db	9037	CGCGCAGAGCTTAAAGTAAACGATCACTGAAGAAATTTCCCACTGAACCAAGAGTTCAAGAA			-----	8978
Qy	1100	ValProGluValThrAlaAspValAspHisValAlaThrCysGln---ValIleLysLeu			-----	1118
Db	8977	ATTATCGAGGAGATCGAAGAGATCGAGGAAGAAAAACCCGACAGAGTATGTGATCCGAAT			-----	8918
Qy	1119	GlnGlnLeuMetGluGlnAlaValAlaProGluSerSerGluThrLeuThrAspSerGlu			-----	1138
Db	8917	ARGAAGCCACACAGAGCTGTA-----GAAGATAGGAAGATTAGCTTACCGAAA			-----	8867
Qy	1139	ThrAsnGlySerThrProLeuAlaAspSerAspThrAlaAspGlyThrGlnGlnAspGlu			-----	1158
Db	8866	AAGAAACCTAAGGCTCCATAGTTGAAGAGCCAGAGCTGAAATAACITTCGAAGCCAAAG			-----	8807
Qy	1159	ThrIleAspSerGlnAspSerLysAlaThrAlaValArgGlnSerGlnValThrGlu			-----	1178
Db	8806	CTAAAGTCGGAAGAAGTTCAGGAAGAAGCAAGATGATGTAAGAAAAACCAAGCAAAATC			-----	8747
Qy	1179	GluGluAlaAlaThrAlaGln-----LysGluGluProSerThrLeuProAsn			-----	1194
Db	8746	GACGAAGTTCGATAGCCGACGAGCTTACGTCAGGTTTGAAGAAGAGGTTGTCCACAG			-----	8687
Qy	1195	AsnValProAlaGlnGluGluHisGlyGlu---GluProGlyArgAspValLeuGluPro			-----	1213

```

Db 8686 CCGATTGTTGAAGAAGATGTAGACAGTTCCGAATTTAAAAAGAAACCAAGACCG 8627
Qy 1214 ThrgInglInIuLeuThraAlaAlaValProValIeu----- 1226
Db 8686 GAGCCAGAGACATGTGTGATCTCGATTTGTGAACCTAAAGAAACAGAGCCAGTAGAT 8567
Qy 1227 ----- 1233
Db 8566 GCTGATGAAGTTGTGCTGAAGTTACCTTAAGCCAAAGCGTAAACCGAATTTACAGAG 8507
Qy 1234 GluGly-----GluValAspTrpLeuAspGlyIuLyValIyGluGly----- 1248
Db 8506 GAAGCATTTCTGTGATGTTAAGCTGCCAAAGAAAGAAAGATTACAGAAATCTTCA 8447
Qy 1249 ---GIngluValPheValHisSerGlyProAsnSerGlnLyAlaAlaAspValThrTyr 1267
Db 8446 GACCAACCTGTACAGCTCAAGAAAAAGAAAGCCGAGAAACCGGTGGAAGAGCGTGGC 8387
Qy 1268 AspSerGluValMetGlyValAlaGlyCysGlnGlyIuLySerThrGluValGlnSer 1287
Db 8386 GACGAGCTTAAAGCTA-----CAGCAAACTGTTGTGAGGAAAGACCT 8345
Qy 1288 LeuSerLeuGluGluGlyGlu---MetGluThrAspValGluLyGluLySerGluThr 1306
Db 8344 GTGGAAATTGAAGAGAAAGAGATCTATTGAGAAAGCACTGTCAATTCGCAAAAGCCAAAG 8285
Qy 1307 LysProGluGlnValSerGluGluGlyGluGluGluThrAlaAlaProGluHisGluGly 1326
Db 8284 AAGCCTTCGAGCGAGCGGTGAGATCTCAAGAAACA-----GAGTTCAAGTCTC 8234
Qy 1327 ThrTyrGlyLysProValLeuThrLeuAspMetProSerSerGluArgGlyLysAlaLeu 1346
Db 8233 TCCTTCAAAAAGCCC---CATACATTAATGAGGTTGAGAGAGCCGCCACAGTTCTG 8177
Qy 1347 GlySerLeuGlyGlySerProSerLeuProAspGlnAspGlyAlaGlyCysTrpLeuVal 1366
Db 8176 AAAAAACGACCGATCAAGCCCTACTCTGATGAAGCTGCAGCT----- 8132
Qy 1367 GlnValGlnSerLeuAspThrThrValThrGlnThrAlaGluAlaValGlu-----Lys 1384
Db 8131 -----GACTTCTTATTAAAGCTCAAGAAAGAAATACGAAGAGGCGAG 8087
Qy 1385 ValIleGluThrValValIleSerGlu-----ThrGly 1395
Db 8086 GATATCGAAGAGCTTCGTGATGTCAGCAAGAAAGCCAGACCTTTGCMAATTCACAGAA 8027
Qy 1396 GluSerProGluCysValGlyAlaHisLeuLeuProAlaGluLysSerSerAlaThrGly 1415
Db 8026 GAGGATGAGGAGGCTTACACAGTGAAGAACTAAAGCGTCGAAAGCA----- 7979
Qy 1416 GlyHisTrpThrLeuGlnHisAlaGluAspThrValProLeuGlyProGluSerGlnAla 1435
Db 7978 -----GTAACAATTCGCAATATGCTGATGTT 7952
Qy 1436 GluSerTrpLeuProIleLeuValThrProAlaProGluSerThrLeuHisProAspLeuGln 1455
Db 7951 GAGAACGTCACATTC-----CGTGCAGAAAGTACCAAGACCAAGGAAGATGTCAT 7901
Qy 1456 GlyGluIleSerAlaSerGlnArgGluAspSerGluGluGly----- 1469
Db 7900 CAAGAAATTCACATTCCTGAGCTGCTATGCCGAGAGAAATTTCCATGCTGGAAG 7841
Qy 1470 -----AspLysPro----- 1478
Db 7840 GTGAAGTTAAAGACCTTACAAAGACCTTCTCGAAGCAGCGGATGAACCAAGATC 7781
Qy 1479 -----AspGlyLysGluSerThrAlaIleGluLysValLeuLysAla 1492
Db 7780 AAGATCATTCAGAGACTTGTGATGATGGTGAAGAACCATTAATCGAAGAAATTCGAGAGAT 7721
Qy 1493 GluProGluIleLeuGluGluGluSerLysSerAsnLysIleValLeuAsnValIleGln 1512
Db 7720 GAGGATACCATTCATGATGATGAGGAAACCAAGAAAGATCTTTGTGAGGAATTTGCCACA 7661

```

```

Qy 1513 ThrAlaValAspGlnPheAlaArgThrGluThrAlaProGluThrHisAlaTyrAspSer 1532
Db 7660 GACCAAGTGAGACTT-CAAGTTTAAGCCAAAGAAAGACCCCAACAGCATATTCGTTC 7602
Qy 1533 Gln 1533
Db 7601 AGA 7599

RESULT 14
ABL28640
ID ABL28640 standard; DNA; 24789 BP.
XX
AC ABL28640;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 37393.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KM pharmaceutical; gene; ds.
OS Drosophila melanogaster.
XX
EN W0200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
RR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE ) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW,
DR WPI; 2001-656860/75.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Claim 1; SEQ ID NO 37393; 21bp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
CC sequences (AB101840-AB116175) and the encoded proteins
CC (AB57737-AB572072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pat_sequences.
XX
SQ Sequence 24789 BP; 7654 A; 5448 C; 6281 G; 5406 T; 0 other;

Alignment Scores:
Pred. No.: 1.17e-17 Length: 24789
Score: 530.50 Matches: 389
Percent Similarity: 36.85% Conservative: 309
Best Local Similarity: 20.54% Mismatches: 639
Query Match: 6.57% Indels: 559
DB: Gaps: 79

US-09-902-432-4 (1-1596) x ABL28640 (1-24789)
Qy 8 GluGlnArgSerProGluGlnProAlaGlySerAspThrProSerGluLeuValLeuSer 27
Db 9664 GACCAAGAGTCC---CAGAAACGAGAAATATCTGAGTCGATGAGCAAGATATCTGAA 9720

```

Qy	28	GlyHisGlyProAlaAlaGluAlaSerGlyAlaAlaGlyAspProAlaAspAlaAspPro	47
		:::	
Db	9721	CGAAAGATTGAAGAACCAAGAAACCTGAAGAAATGGACACTGAAGCTAAATATCAGAGAA	9780
Qy	48	AlaThrLysLeuProGlnLysAsnGlyGlnLeuSerSerValAsnGlyValAlaGlu---	66
		:::	
Db	9781	GCTACAGCTTGGATAGCAAGTCTCGGAGAAAGAAAGAAAGAACTAGAGGCTAGTGCAGAAAG	9840
Qy	67	GlnGlyAspValHisValGlnGlnGluAsnGlnGlnGlyGlnGlnGluGluGluValValAsp	86
		:::	
Db	9841	CAAGGTGATCAGGATGTCCAAAGAAAGTCCCGAGAAACCCAGAAAGTATCTCTGAGGTCGTAGCT	9900
Qy	87	GluAspValGlyGlnArGluSerGluAspValArGlu-----	99
		:::	
Db	9901	GAGAAGATATCTGAAGAAACAATTTGAAGAGCCAAAGAAACCTGGAAGTAAAGGACACTGAG	9960
Qy	100	-----	105
		:::	
Db	9961	ATTAAATCAGAGAAAGCCACAGCCTTGGATTAAGCAAGTCTTGGAAAGAAAGAAAGAACTAGAG	10020
Qy	106	MetAlaAlaAsnSerThrAlaValGluAspIleThrLysAspGlyGlnGlu---GluThr	124
		:::	
Db	10021	GCTAGTCACAAAAGCAATGTGATCAGGATGTTCGAAGAAGAGTCCCGAGAAACCCAGNAGTA	10080
Qy	125	SerGluIleIleGluGlnIleProAlaSerGluGlnAsnValGluGluMetValGlnPro	144
		:::	
Db	10081	TCTGAGGTCGTAGCTGAG---AAAATATCTGAAGAAACCAATTTGAAGAGCCAAAGAAACCT	10137
Qy	144	-----	144
Db	10138	GAAGTAAAGGAAACCTGAGGTTAAATCAGAGAAAGCTACAGTCTTGGATAAGCAAGTCTCTG	10197
Qy	145	-----AlaGluSerGlnAlaAsnAspValGly-----	153
		:::	
Db	10198	GAAGAAAGGAACTAGAGCTAGTGCCACAAAGCAAGGTGATCAGGATGTCGAAAGNAG	10257
Qy	154	PhelysLys-----ValPheLysPheValGlyPheLysPheThrValLys-----	168
		:::	
Db	10258	TTCCAGAAAGCAGAAAGTATCTGAGTCGTGAGAGAGATATCTGAAAGAAACAAATTGAA	10317
Qy	169	-----LysAsp-----LysAsnGluLysSerAspThrVal	178
		:::	
Db	10318	GAGCCAAAGAAACCTGAAGTAAAGGACACTCAGATTAATCAGAGAAAGCCACAGCCTTG	10377
Qy	179	GlnLeuLeuThrValLysLysAspGluGlyGlnGlyAlaGluAlaSerValGlyAlaGly	198
		:::	
Db	10378	GATAGCAAGTCTCGAAGAAAGAACTAGAGGCTAGTCGACAAAGCAAGGTGATCAG	10437
Qy	199	Asp-----HisGlnGluProSerValGluThrAlaValGlyGluSerAlaSer	214
		:::	
Db	10438	GATGTCGAAAGAAAGTCCCGAGAAACCAAGATATCTGAAAGTATCAGTGTGAGAGATATCT	10497
Qy	215	LysGluSer---GluLeuLysGlnSerThrGlnLysGlnGlnGlyThrLysLysGlnGlu	233
		:::	
Db	10498	GAGGAAAGATTTGAAGAACCAAGAAACCTGAAGAAAGGAAAGAACTGAGGTTAAATCAGAG	10557
Qy	234	GlnSerSerThruIleProLeuGlnAlaGluSerAspGlnAlaAlaGluGluGluAla	253
		:::	
Db	10558	AAA---GCTACAGTCTTGATAGCAAGTCTCTGGAAAGAAAGGAAAGAACTAGAGGCTAGTGCA	10614
Qy	254	LysAspGluGlyGlu-----GluLysGlnGluLysGluProThrLysSerPro---	269
		:::	
Db	10615	CAAAAGCAAGGTGATCAGGATGTCGAGAAAGGTCCTCCAGAAACCCAGAGATATCTCAGGTC	10674
Qy	270	-----GluSerProSerSerProValAsnSerGlu	279
		:::	
Db	10675	GTAGCTGAGAAGGTATCTCAAGGAAAGATTTGAAGAAACCAAGAAACCTGGAAGTAAAGGAA	10734
Qy	280	ThrThrSerSerPheLysLysPhePheThrHisGlyTTPAlaGly---TrpArgLysLys	298
		:::	
Db	10735	ACAGAGCTAAATCAGAGAAAGCCCAACTTT-GGATATGCAAGTCTCTGGAAAGAAAGAGA	10793
Qy	299	Thr-SerPheLysLysSerLysGluAspLeuGluThrAlaGluLysArGlyLysGluGlu	318

Db	10794	ACTAGAGGCTAGTGCACAAAAGCAAGGTGATCAGGATCTCGAAAAGAAGTCCCGAGAAACC	108513
Qy	318	nglu-----AlaGluLysValAspGluGlu-----	326
Db	10854	AGAGTATCTGAAGTCATAGCTGAGAGAGATATCTGAGAAAAGATTGAAGAACCAAGAA	10913
Qy	327	----GluLysGluLysThrGluProAlaSerGlu-----	336
Db	10914	ACCTGAAGAAAGGAAACCTGAGGTTAAATCAGAGAAAGCTACAGTCTTTGGATAAGCAAGT	10973
Qy	337	----GluGlnGluProAlaGluAspThrAspGlnAlaArgLeuSerAlaAspTyrGlu--	354
Db	10974	CCTGAAGAAAGAAAGAACTAGAGGCTAGTGCACAAAGCAAGGTCATCAGGATGTCGAAAA	11033
Qy	355	-----LysVa	356
Db	11034	GAAGTCCAGAAAACCAAGAGTATCTGAGGTCGTAGCTGAGAGGTTATCTGAAGGNAAGAT	11093
Qy	356	IgLeuLeuProLeuGluAspGlnValGlyAspLeuGluAlaSerSerGluGluLysCysAl	376
Db	11094	TGAAGAACCAAGAAACCTGGAAGTAAAGGAAACCTGAGGCTTAAATCA--GAGAAAGCCAC	11150
Qy	376	aProLeuAlaThrGluValPheAspGluLysMetGluAlaHisGlnGluValValAlaG1	396
Db	11151	AACTTTGGATAGCAAGTCCTGGAAGAAAA--	11181
Qy	396	uValHisValSerThrValGluLysThrGluGluGluGlnGlyGly--GlyGlyGluAl	415
Db	11182	-----GAAGTAGAGGCTAGTGCACAAAGCAAGGTATCAGGATGGAAAAATC	11288
Qy	415	aGluGlyGlyValVal-----ValGluGlyThrG1	425
Db	11229	TCGAGATGATATTATAAAACATTCAAGAAAGGCTCAGAGATTTATCCAGGCCATTAGG	11288
Qy	425	yGluSerLeuProProGluLysLeuAlaGluProGlnGluValProGlnGluAlaGluPr	445
Db	11289	ATCTAGTGTG--GATGAAATTTTGAGAGAAAGTCGGGAAATTTGTCAATACCTTTGAAGA	11345
Qy	445	o-----AlaGluGluLeuMetLysSerArgGluMetCysValSerGlyGlyAs	461
Db	11346	TGATAGGTAGTGGCAAAACATTTATTCAAATTAAGAGACCATATAGTGTACACCTATGA	11405
Qy	461	phisThrGln-----Le	465
Db	11406	CGAARAAAGAGAGAGAAGAAATAAGAAAAGAACTCTTCGAACTCTTTATTGAATGTT	11465
Qy	465	uThrAspLeuSerProGlu-----	471
Db	11466	GTGTGAGGCCTCCCTGGAAGCAGCTGAGAAAGTAAACTCAACTATCTTAAGGAAATTA	11525
Qy	471	-----	471
Db	11526	AACCAATGTGATTTTAAACAAAGCTACTATTCAACTTATCGACGACTCTACATGTTCCAC	11585
Qy	471	-----	471
Db	11586	CAACCGTCCTTATTAAATCCCAAACTTTTAAATTTTGAGAGAGTTCGCGTTAAATTC	11645
Qy	472	----GluLysThrLeuProLysHisProGluGlyLeValSer-----	484
Db	11646	ATCAGAAACATATGTTGATAAGTCATCGGAAAAAATGATTAGTTTGCACACAGAGCTTAT	11705
Qy	484	-----	484
Db	11706	GGATATTTTCGTAATCTTTGTATGACTTTTATAGATCAGAGACGGAAGTTTTTAAACCAAA	11765
Qy	485	-----GluValGluMetLeuSerSerGlnGluArgIleLysValGlnGlySe	500
Db	11766	AATCGAAAAATATAAACACAGCTGCTCAGTGACTATGATTATATTGAAAAAAGATCG	11825
Qy	500	rProLeuLysLysLeuPheSerSerSerGlyLeuLysLysLeuSerGlyLysLysGlnLy	520

Db 11826 ACCCTCTTAACGCTGTAATATGAAAGATTAATGTTATCC----- 11871
 Qy 520 sGlySarArgIyGlyGlyAspGluGluProGlyGluIleHisThr-- 539
 Db 11872 -----CAACATATCTTAATCTAT 11888
 Qy 540 -----GluSerProGluSerAlaAspGlu---GlnYsGlyGluSerSe 553
 Db 11889 TATTGAGAAAGTTAAACATGATGACTGAATAATCATATGATCAAAAGAAAAAGACGTGCTCAA 11948
 Qy 553 rAlaSerSerProGluGluGluProGluGluThrCysLeuGluYsGlyProLeuGluAl 573
 Db 11949 CCGGAGCGCGGTAATTTCCGCCGACGAAAAAGCA-----GAAGAAAGTCAAAAGAGAGA 12002
 Qy 573 aProGlnAspGlyGluAlaGluGluGlyThrThrSerAspGlyGluYsYsArgGluG 593
 Db 12003 GATTAAAGACCTCTGAAGCAACACAGAAATCGAAAGTTTCGAGAGAAATCA----- 12057
 Qy 593 yIleThrProTrpAlaSerPheYsLeuYsMetValThrProYsYsArgValArgArgPr 613
 Db 12058 -----ATCGAAGAAAGAACTTGAAAGCAAGAAAGAAACAGACAGAGTC 12104
 Qy 613 oSerGluSerAspYsGluGluGluGluYsValYsSerAlaThrLeuSerSerTh 633
 Db 12105 TGCTATTGATGAAAAATCCCAAAAGGCTGAGGTGCTCCGAAACTGCTCCGAAAAAGATTAC 12164
 Qy 633 rAspSerThrValSerGluMetGln---AspGluValYsThrValGlyGluGluGly 652
 Db 12165 TCATGAGAAAGCAACAAAGAAAGCAAAAGAGAGAGGTTAAG-----GATTTCGAAGCCAA 12218
 Qy 652 sProGluGluProYsArgArgValAspThrSerValSerTrpGluAlaLeuIleCysVa 672
 Db 12219 GCCTTAAGAAAGCAAGGCTTGTGAGAAAGAAATCAATCGAAGAAAGAAACTT----- 12270
 Qy 672 lGlySerSerYsYsArgAlaArgYsAlaSerSerSerSerAspAspGluGlyGlyProAr 692
 Db 12271 -----GAAGCAAGAAAGAAACAGACGAGTGTGCTATTGATGAA----- 12312
 Qy 692 gThrLeuGlyGlyAspSerHisArgAlaGluGluAlaSerYsAspYsGluAlaGlyTh 712
 Db 12313 -----AAATCCCAAAAGCTGAGGTGCTCGAAATTGCTCCGAAAAAGATTAC 12359
 Qy 712 rAspAlaValProAlaSerThrGlnGluGlnAspGlnAlaGlnGlySerSerSerProG 732
 Db 12360 TCATGAGAAAGCAACAAAGAAAGCAAAAGAGGTTAAGATTCTGAAGCCAAAGCTTAA 12419
 Qy 732 uProAlaGlySerProSerGluGlyGluGlyValSerThrTrpGluSerPheYsArgHe 752
 Db 12420 GAAGCA-----AAGGT 12431
 Qy 752 uValThrProArgYsYsSerYsSerYsLeuGluGluYsAlaGlu-----As 769
 Db 12432 CTTGAGAAAGAAATCAATCGAAGAAAGAAACTTGAAAGCAAGAAAGAAACAGACGGA 12491
 Qy 769 pSerSerValGluGlnLeuSerThrGluIleGluProSerArgGluGluSerTrpValSe 789
 Db 12492 CTCTGCTATTGATGAATAATCCCAAAAGGCTGAGTCCGAAATTGCTCC----- 12543
 Qy 789 rIleYsYsPheIleProGlyArgArgYsYsArgAlaAspGlyYsGlnGluGlnAl 809
 Db 12544 -----GAAAAATTACTGATGAGAAAGCAACAAAG 12575
 Qy 809 aThrValGluAspSerGlyProValGluIleAsnGluAspAspProAsnValProAlaVa 829
 Db 12576 TCAAAAGAGAGAG-----CTTAAGATTCTGAAGCCAAAGCTTAAGAGCCAAAGGT 12626
 Qy 829 lValProLeuSerGluYsAsnAlaValGluArgGluYsMetGluAlaGlnGlyAsnTh 849
 Db 12627 CTTG-----GAGAAAGAAATCAATCGAAGAAAGAAACTTGAAAGCAAGAAAGAAAA 12677
 Qy 849 rGluLeuProGlnLeuLeuGlyValAlaYrValSerGluGluGluSerYsYsThrLeuY 869
 Db 12678 ACAGACGAGCTGCTGTAAT-----GATGAAAAATCCCAAAAGGCTG- 12718

Qy 869 aHisThrValSerValAlaValIleAspGlyThrArgAlaValThrSerValGluGluA 889
 Db 12719 -----A 12719
 Qy 889 rGserProSerTrpIleSerAlaSerValThr-----GluProL 902
 Db 12720 GGTGCTCGAAATGTCTCCAAAAAGATTACTGATGAGAGGCAACAAAGAAAGCCAAAAAGA 12779
 Qy 902 euGlnHisThrAlaGlyGluAlaMetProProValGluGluValThrGluYsAspIleI 922
 Db 12780 AGGAGGTTAAGGTTCTGTAAGCCAAAGCTTAAGAAAGCAAGGCTTGGAAGAAATCAATCA 12839
 Qy 922 lValGluGluThrProValLeuThrGlnThrLeuProGluGlyYsAspAlaHisAspA 942
 Db 12840 TCGAAGAAAG-----AAACTTGAAACAAAGAAAGAAACAGACAGCG 12881
 Qy 942 sMetValThrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThrSerG 962
 Db 12882 AC-----TCTGCTATTGAT-----GAAAAATCC 12905
 Qy 962 lu-AlaLeuArgThrGluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMet 981
 Db 12906 AAAAGCTTGAGGTGCTCGAAATTTGTC-----TCGAAAGATTAAGTATGAG 12953
 Qy 982 ValSerAlaValSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProVal 1001
 Db 12954 AAGGCAAGAAAGCAAAAGAGAGGTTAAGATTCTGAAGCCCTTAAGAAACA 13013
 Qy 1002 GlnGluValGluSerGlyValLeuAspThrGlu-----GluGluGluArgGlnThrGln 1019
 Db 13014 AAGGCTTGAGAGAAATCAATCAGAAAGAAAGAACTTGAAAGCAAGAAAGAAACACAG 13073
 Qy 1020 AlaIleLeuGlnAlaValAlaAspYsValYsGluGluSerGlnValProAlaThrGln 1039
 Db 13074 ACG-----GACTCTGCTATTGATTAATAATCCCAAAAGCTGAGGTGCTCGAAATTTGCT 13127
 Qy 1040 ThrValGlnArgThrGlySerYsAlaLeuGluYsValGluGluValGluGluAspSer 1059
 Db 13128 CCGAAAGAAATTACTTGATGAGAGGCAACAAAGTCAAAAGAGAGAGGTTAAGATTCT 13187
 Qy 1060 GluValLeuAlaSerGluYsGluYsAspYsAlaMetProYsGlyProValGlnGluAla 1079
 Db 13188 GAA-----GCCAAGCTTAAGAAAGCAAAAGCTTGAGAAAGAAATCAATCGAAGAACG 13241
 Qy 1080 GluAlaGlnHisLeuAlaGlnGlySerGluThrGlnAlaThrProGluSerLeuGlu 1099
 Db 13242 AAATCTGAAGCAAGAAAGAA-----ACACAGAGGAGTGTGCTATTGAGAAATCCCAA 13298
 Qy 1100 ValProGluValThrAlaAspYsAlaAspHisValAlaThrCysGlnValIleYsLeuGln 1119
 Db 13299 AAGGCTGAGGTGCTCGAAATTTGCTCCGAAAAAGATTACTGATGAG-----AAGGCACAA 13352
 Qy 1120 GlnLeuMetGluGlnAlaVal-----AlaProGluSerSerGluThrLeu 1134
 Db 13353 GAAAGCCAAAGAAAGAGGTTAAGGTTCTTAAGCCAAAGCTTAAGAAAGCAAGAGTCTTG 13412
 Qy 1135 ThrAspSerGluThrAsnGlySerThrProLeuAlaAspSerAspThrAlaAspGlyThr 1154
 Db 13413 -----GAGAA-AGAAATCAATCGAAGAAAGAAACTTGAAACCAAGAAAGAAAAA 13462
 Qy 1155 GlnGlnAspGluThrIleAspSerGlnAspSerYsAla---ThrAlaAlaValArgG 1173
 Db 13463 CAGACGAGAGTGTGCTATTGATGAAATAATCCCAAAAGGCTTGAGGTGCTCGAAATTTGCTC 13522
 Qy 1173 nSerGlnValThrGluGluGluAlaAlaThrAlaGlnYsGluGlu-----ProSe 1190
 Db 13523 CGAAAGATTAAGTGAAGAGGCAAGAAAGCCCAAGAAAGAGGTTAAGATTCTGA 13582
 Qy 1190 rThrLeuProAsnAsnValProAlaGlnGluGlnHisGlyGluGluProGlyArgAspYsVa 1210
 Db 13583 ACCCAAGCTTAAGAGGCAAGAGCTTGTGAGAAAGAAATCAATCGAA-----GAGAGAGA 13636

QY 1210 lLeuGluProThrGlnGlnGluLeuThrAlaAlaValProValLeuAlaLysThrGI 1230
 Db 13637 ACTTGAAACAGAGAAAGAAACAGACGGAGTCTGCTATT-----GATGA 13681
 QY 1230 uValGlyGlnGluGluValAspTrpLeuAspGlyGluLysValLysGluGluGlnGI 1250
 Db 13682 AAAATCCAAAGGCTGAGGTGCCGAACTGTCTCCGAAAGATTACTGATGAGAAGGC 13741
 QY 1250 uValPheValHisSerGlyProAsnSerGlnLysAlaAlaAspValThrTyrAspSerGI 1270
 Db 13742 A-----GAAGAAAGCCGAAAGGAGGAGGTAAAGGACTCTGA 13777
 QY 1270 uValMetGlyValAlaGlyCysGlnGluLysGluSerThrGluValGlnSerLeuSerLe 1290
 Db 13778 A-----GCCAAGCTTAAGAAAGCAAGGCTCTTGAGAAAGAAATCAAT 13819
 QY 1290 uGluGluGlyGluMetGluThrAspValGluLys----- 1301
 Db 13820 CGAAGAGAGAACTTGAACACAGAAAGAAACACACAGAGTCTCTATTGATGAAA 13879
 QY 1302 -----GluLysArgGluThrLysProGluGlnValSerGluGluGluGI 1317
 Db 13880 ATCCCAAAAGGCTGAGGTGCCGAACTGTCTCCGAAAGATTACTGATGAGAAGGCACA 13939
 QY 1317 nGluThrAlaAlaProGluHisGluGlyThrTyrGlyLysProValLeuThrLeuAspMe 1337
 Db 13940 AGAAAGCCAAAGAGAGGTTAAGGATTCTGAAGCCAGCCT----- 13982
 QY 1337 tProSerSerGluArgGlyLysAlaLeuGlySerLeuGlySerProSerLeuProAs 1357
 Db 13982 ----- 13982
 QY 1357 pGlnAspLysAlaGlyCysLeGlu-----ValGlnValGlnSerLeuAspThrTh 1374
 Db 13983 -----AAGAAAGCAAGATTTTGGAGAGAAATCAATCGAAATAGAGAAATTTGGAT----- 14033
 QY 1374 rValThrGlnThrAlaGluAlaValGluLysValIleGluThrValValIleSerGluTh 1394
 Db 14034 -----GAAAGAAAGAAAGAAAGCAGCAGGAGCTAAGGTTGCCACTGATAC 14077
 QY 1394 rGlyGluSerProGluCysValGlyAlaHisLeuLeuProAlaGluLysSerSerAlaTh 1414
 Db 14078 C-----AAATCCCAACGGTTGAGTTTCCGAAATCGTTTGGAAAGATATCGGAA-- 14129
 QY 1414 rGlyGlyHisTrpThrLeuGlnHisAlaGluAspThr-----ValProLeuGlyProGI 1432
 Db 14130 -----CAAAAGGCAGAGAAAGCAAAAGCTTGAACTT---AAGGA 14167
 QY 1432 uSerGlnAlaGluSerIleProIleValThrProAlaProGluSerThrLeuHisPr 1452
 Db 14168 CTCTGAGGCCAAATCAAAAGCAAGCAAGGTCCTG---GAAAAGAAAGTCAACCTTGAAAGA 14224
 QY 1452 oAspLeuGlnGlyGluIleSerAlaSerGlnArgGlu-----ArgSerGI 1467
 Db 14225 GAACCTCGAT---GAGATGATAAAACAGAGAGGAGATGGCGCTACTAACAAATCCCA 14281
 QY 1467 uGluGluAspLysProAspAlaGlyProAspAlaAspGlyLysGluSerThrAla----- 1485
 Db 14282 AAAGCGAGGCTGCTGATGTTGTTCCGAAAGAAATATCTGAAGAGAGGTTGCTGAAT 14341
 QY 1486 -----IleGluLysValLeuLysAlaGluProGluIleLeuGluGI 1500
 Db 14342 AAAAAGCTCTGAACCTATGAGCTCAAGGCTAAATCTAAACACAGATGATTACTGCGCA 14401
 QY 1500 uSerLysSerAsnLysIleValLeuAsnValIleGlnThrAlaValAspGlnPheAlaAr 1520
 Db 14402 TGAAGAGTACAGGGGCTAAGGTTCTCGAATCTGTTCTGTTAAAGAAATGAAGCTGAGAA 14461
 QY 1520 gThrGlu-----ThrAlaProGluThrHisAlaTyrAspSerGlnThrGlnValProAl 1538
 Db 14462 AACAGATCAACTTTCAGCAAGAAACCAACTGTTCTGGATGAGGATCTGTTGTTCCAAA 14521
 QY 1538 aCysArg----- 1540

Db 14522 ACAGAAAGCCATATTGGCTGAACAAATCGAGATTTATTAGCTGCAACAGTACAAGTC 14581
 QY 1541 -LeuAspSerArgGluProAsnArgCysTrpThrLysMetLysAspAlaLysMetLysHi 1560
 Db 14582 CATGGACTCCGAGTACAGGACCGA-----AAGGAATCCGAGAGTGCACAGCGGAAG-- 14633
 QY 1560 sProValProGlnProArgGluAspLeuGlnVal 1571
 Db 14634 -----CCACCGTTGACATCCAACCTA 14654

RESULT 15

AAS80823 standard; cDNA; 7568 BP.

XX AAS80823;

XX 13-FEB-2002 (first entry)

XX DNA encoding novel human diagnostic protein #16627.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI: 2001-639362/73.

XX P-PSDB; ABG16636.

XX New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity -

XX Claim 1; SEQ ID No 16627; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention.

XX Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 7568 BP; 2312 A; 1873 C; 1825 G; 1538 T; 0 other;

Alignment Scores:

Pred. No.:	3,286-16	Length:	7568
Score:	492.00	Matches:	386
Percent Similarity:	33.47%	Conservative:	269
Best Local Similarity:	19.72%	Mismatches:	714
Query Match:	6.09%	Indels:	589
DB:	23	Gaps:	75

US-09-902-432-4 (1-1596) x AAS80823 (1-7568)

```

QY 45 AlaAspProAlaThrLysLeu-----ProGlnLysAsnGlnGlnLeuSer 59
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1593 GCAAAACCTGCGGAGAAATCATCCAGTCTGTTCTCTGGAAACGACACCAAGTCAAC 1652
QY 60 SerValAsnGlyValAlaGlnGlnGlnLysAspValHis-----ValGln 73
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1653 ATCTCGAAGAGGTTGGAAAGCTCAAAACATCTAGACTTCTGAAGCAGCCACTGGCACCC 1712
QY 74 GlnGlnLysAsnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 93
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1713 CAAAAGGATCTCATCTGCGAGGCTCCCTCTGTGTGAAACAAACAACTGAAGACG 1772
QY 94 SerGlnAspValArgGlnLysAspArgValGlnGlnMetAlaAlaAsnSerThrAlaVal 113
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1773 AGGGCTGATAGCCGAGAAAGT-----CTGAAGCCAGCCGCAAAACCACTCTTAC 1823
QY 114 GlnAspIleThrLysAspGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 133
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1824 AATCCGTGCGAGGAGTCAAAAGAAAGAACCCCTGAGAGTC-----ACA 1868
QY 134 SerGlnAsnAsnValGlnGlnMetValGlnProAlaGlnSerGlnAlaAsnAspValGly 153
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1869 AAAGTCAATCAGCTGGAA-----AAGCCACCCAAAGTTGAAAGCAAGAA----- 1913
QY 154 PheLysLysValPheLysPheValGlnPheLysPheThrValLysLysAspLysAsnGln 173
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1914 -----AAGTAAATGTTGAAAGAAAGCAAGCAAGCAAGTAA 1943
QY 174 LysSerAspThrValGlnLeuLeuThrValLysLysAspGlnGlnGlnGlnGln 193
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1944 AAAAAAGAGACCAAAACCTTCAGTACTGAAAGAGAG----- 1979
QY 194 SerValGlyAlaGlyAspHisGlnGlnGlnProSer--ValGlnThrAlaValGlyGlnSer 212
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1980 -----GTTCCCAAGCAAGAGAGCCATCTCCAGTAAAGCCGAGTGGCTGAG-- 2027
QY 213 AlaSerLysGlnSerGlnLeuLysGlnSerThrGlnLysGlnGlnGlnGlnGln 232
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2028 -----AAGCAAGCCACAGATGTCAAACCCAAAGCTGCCAAGAG 2066
QY 233 GlnGlnSerSerThrGlnIleProLeuGlnAlaGlnSerAspGlnAlaAlaGlnGlnGln 252
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2067 AAGACCGGTAAAAAGAGAAACAAAGGTAAAGCCTGAAGACAAAGAAAGAGAGAAAG 2126
QY 253 AlaLysAspGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 272
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2127 CCAAGAGAAAGAGTGGCTAAAAAGAGAGCAAAACCACTATCAAGAGAGAGAAAGCA 2186
QY 273 SerSerProValAsnSerGlnThrThrSerSerPheLysLysPheSerThrHisGlyTyr 292
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2187 AAA-----AAGGAAGAGGTGAAAAAGAAAGTCAAAAAAGAGATC----- 2225
QY 293 AlaGlyTyrArgLysLysThrSerPheLysLysSerLysGlnAspAspLeuGlnThrAla 312
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2226 -----AAGAAAGAGAGAGAAAGAAAGAAAGAGAGAGTAAAGAAAGAAAGAGAG 2276
QY 313 GlnLysArgLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 332
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2277 CCAAGAGAGAGTCAAGAGAGAGTAAAGAG-----GAAGAGAGAGAGAGAG 2321
QY 333 ProAlaSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 352
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2322 GTGAAAAAGAGAAAGAGAAAGCAAAAGAAAGATTT-----AAGAGAGCTCCTTAAAGAC 2375

```

```

QY 353 TyrGlnLysValGlnLeuProLeuGlnAspGlnValGlyAspLeuGlnAlaSerSerGln 372
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2376 GCAAGAAATCATCTACCTCTCTG-----TCTTAAGCA 2408
QY 373 GlnLysCysAlaProLeuAlaThrGlnValPheAspGlnLysMetGlnAlaHisGlnGln 392
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2409 AAAAAACAGCTGCTTTAAAAACCAAAAGTACCAGAGAGAGAGAGTCTGTACAGAAAGAT 2468
QY 393 ValValAlaGlnValHisValSerThrValGlnLysThrGlnGlnGlnGlnGlnGln 412
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2469 TCTGTGCTGCCGGAAGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2528
QY 413 GlnGlnAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 432
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2529 AAGGCGGAGAGGCTGTGCTGCAAGCTGTGCGACATGTGAGGACCAAGCAAGCTGTGATG 2588
QY 433 LeuAlaGlnProGlnGlnGlnValProGlnGlnAlaGlnProAlaGlnGlnGlnMetLysSer 452
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2589 GCGGAGCTGGAATAGCAAGC-----ATTGACCTGCGCAAGAAAGTGAAGCTGAG 2639
QY 453 ArgGlnMetCysValSerGlyGlyAspHisThrGln--LeuThrAspLeuSerProGln 471
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2640 AGGTCCCTTATGTATCTCTGAGAGATCTAACCAAGACTTTGAAGAGTTAAAGGCTGA 2699
QY 472 GlnLysThrLeuProLysHisProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 491
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2700 GAGTCGATGTAAACAAAG-----GACATCAAGCTCTGAGCTGAGCTATCAAGAC 2750
QY 492 GlnGlnLysGlyLeuValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 511
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2751 GAAGAGAAACGTGAAG--GAATCTGAGCCAGTCAAGCTTCATCATCAAGAGAGAGA 2807
QY 512 LysLysLeuSerGlyLysLysGlnLysGlnLysLysArgGly----- 524
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2808 GAAGTCACCAAGAGTCCGCGAGTCCCTGATGAGGAATCACTACCACTGAAGGGAG 2867
QY 555 GlyGlyLysAspGlnGlnGlnProGlyGlnGlnGlnGlnGlnGlnGlnGlnGln 544
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2868 GCGGAGATGTGAACACACCTGAGAGGCTGAGCCCTGCGAAGACAGAGATAC-- 2924
QY 545 AlaAspGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 564
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2925 --GACATTGAAAAATTTGAAGATCAAGAGAGCGGTTTGAAGATCTTCAAGAGCT-- 2978
QY 565 CysLeuGlnLysGlyProLeuGlnAlaProGlnAspGlyAlaGlnGlnGlnGlnGln 584
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2979 -----GAGACTATGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3026
QY 585 SerAspGlyGlnLysLysArgGlnGlyIleThrProThrAlaSerPheLysLysMetVal 604
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3027 GAGGATGGGAGAGACAC-----GTA 3047
QY 605 ThrProLysLysArgValArgArgProSerGlnSerAspLysGlnGlnGlnGlnGln 623
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3048 TGTGTGAGCGCTCCAGAGACAGCCCACTGAGATGAGAGAAAGTGCAGAGCGGAGGCT 3107
QY 624 -----LysValLysSerAlaThrLeuSerSerThrAspSerThrValSerGln 639
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3108 GATGATACATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3167
QY 640 MetGlnAspGlnValLysThrValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 659
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3168 GACATGATGAGGAGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3227
QY 660 ValAspThrSer-----ValSerTyrGlnAlaLeu 669
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3228 GAGGACAAAGCTGAAGATGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3287
QY 670 IleCysValGlySerSerLysArgLysArgLysAlaArgLysAlaSerSerSerAspAspGln 689
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3288 GACTATGTGATGCTGTGTGTGACAGAGCTGCAAGAGCTGTGTGTGTGCGAGAGAGAGAG 3347

```

Qy	690	Gly	-----ProArgThrLeuGlyGlyAspSerHisArgAlaGluGluAla	704
Db	3348	GGATTCCTCCACCAACCAAGCAACTAGGAGCCCACTCTCTCTGGCCGAGAACTTCGA	3407	
Qy	705	SerLysAspLysGluAlaGlyThrAspAlaValProAlaSerThrGlnGlnAspGln	724	
Db	3408	TCITCAATTCATGAT-----GAGACTTTACCTGGAGGCTCAGAGCGGAGGCCACC	3458	
Qy	725	AlaGlnGlySerSerProGluProAlaGlySerProSerGluGlyGluGlyValSer	744	
Db	3459	GCTTCTCATCAGGAGAATCGAGAA-----GACCAAGCCTGAGGAATTCACCTGCCACCTCT	3512	
Qy	745	ThrTrpGluSerPheLysArgLeuValThrProArgLysLysSerLysSerLysLeuGlu	764	
Db	3513	GGCTAC-----ACTCAGTCTACTATTGAG	3536	
Qy	765	GluLysAlaGluAspSerSerValGluGlnLeuSerThr-----	777	
Db	3537	ATATCCAGTCAGCCACCCCTGGATGAGATGTCACCTCGACGCTGATGATGAT	3596	
Qy	778	-----GluLeuGluProSerArgGluGluSerTrpValSerIleLysLys---	792	
Db	3597	GAGACCAACAATGAAGAGCGGAGTCCCTCTCTCAGGAATTCGTAATAATATCACCAATAT	3656	
Qy	792	-----	792	
Db	3657	GAATCTTCATTGTATTCTCAGGAATACTCTAAACCTGCTGATGTTACACCGCTCAACGGA	3716	
Qy	793	PheIleProGlyArgArgLysLysArgAlaAspGlyLysGlnGluGlnAlaThrValGlu	812	
Db	3717	TTTTCTGAAGATCAAAAACAGATGCCACTGATGCAAGGATTACAATGCTTCAGCCTCT	3776	
Qy	813	AspSerGlyProValGluIleAsnGluAspAspProAsnValProAlaValProLeu	832	
Db	3777	ACCATATCACACCC-----	3791	
Qy	833	SerGluTrpAsnAlaValGluArgGluLysMetGluAlaGlnGlyAsnThrGluLeuPro	852	
Db	3792	-----TCITCCATGGAGGAACAATAATTCAGCATCTGCT-----	3827	
Qy	853	GlnLeuLeuGlyAlaValTyrValSerGluGluLeuSerLysThrLeuValHisThrVal	872	
Db	3828	-----TTAGCTGATGCTTACTGCTCTGAAGTGAAGCCAGCACCC-----	3866	
Qy	873	SerValAlaValIleAspGlyThrArgAlaValThrSerValGlu-----GluArg	889	
Db	3867	ACTTTGGACATCAAGATAGCATCTCAGCTGTTTCAAGTGAAGAGTCAGCCCATCGAAG	3926	
Qy	890	SerProSerTrpIleSerAlaSerValThrGluProLeuGluHisThr---AlaGlyGlu	908	
Db	3927	AGCCCGTCC---CTGAGTCCATCTCCACCATCACCTTTAGAAAAGACCCCTGGGTAA	3983	
Qy	909	-----AlaMetProValGlu---GluValThrGluLysAspIleIle	922	
Db	3984	CGTAGTGTGAACCTCTCTCTGACGCCAATGAGATTAAGTCTCTGCAGAGCGCAAGATA	4043	
Qy	923	AlaGluGluThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHisAspAsp	942	
Db	4044	GCCCCGGTGTCTCTCAGGTGACCCCAAGAAGTAGTTGAAGAACATTGTGCTAGTCTCTGAG	4103	
Qy	943	MetValThrSerGluValAspPheThrSerGluAlaValThrAlaThr-----	958	
Db	4104	GACAAGACTCTGGAAGTGGTGTCCACCATCTCAGTCCGCTGAGTGGCAGTGTGGTGCACACA	4163	
Qy	959	-----GluThrSerGluAlaLeuArgThrGluGluVal---	969	
Db	4164	CCTTACTATCAATCTCTACTGACGAGAAATCCAGTCATCTCCCTCAGAGAATCATTTGAA	4223	
Qy	969	-----	969	
Db	4224	AAACCCACGACGTTCCAGTGAGTTTGAATTCAGTGATGCCAAGATGAGATGAAGG	4283	
Qy	969	-----	969	

Db	4284	CGTTCAGTAAGACCCCAATGGATGAGCCGCTGCTGACTCAGAGTCTCTCTATTGAAAGAGTT	4343
Qy	969	-----	969
Db	4344	TTGTCTCTTTACGACGCCGCCCTCATTTGGATCCGAGTCTGTCTATTGAAAGTTTTCTA	4403
Qy	970	-----	974
Db	4404	AGTCTGATGACACAAGCTTTCTGGCAGAGGTGCCGAAAGTCTCTTTTGAAGAAAGAGTGGGA	4463
Qy	975	AlaGluGluThrThrAspMetValSerAlaValSerGlnLeuThrAspSer-	991
Db	4464	AAACAAGGCTCTCCAGACCAAGTAAGTCCAGTTTCTGAAATGACTTCTACTAGTCTTTAC	4523
Qy	992	ProAspThrThrGlu-----GluAlaThrProValGlnGluValGluSerGly	1007
Db	4524	CAAGACAAACGAGGAGGGAACACAGACTTTGCACCAATAAAGAAGACTTTGGCCAA	4583
Qy	1008	ValLeuAspThrGluGluGluArgGlnThrGlnAlaLeuGlnAlaValAlaAsp	1027
Db	4584	GAAGAAGAAACTGATGATGTTGAAGCCATGAGTTCTCAACGAGCACTGGGCTCTGGTAGAA	4643
Qy	1028	Lys-----	1028
Db	4644	AGGAAATTAGGAGATGTTTCTCCACACAAATAGATGTCAGTCAGTTGGATCTTTTAAA	4703
Qy	1029	-----	1037
Db	4704	GAAGACACTAAGATGTCCTATTTCTGAAGTACTGTCTCAGACCAAGTCAGTACTCTCTGT	4763
Qy	1038	ThrGlnThrValGlnArgThrGlySerLysAlaLeuGluLysValGluValGluGlu	1057
Db	4764	GATGAGGCGGTACACAGACAGACGACTCTCATATGGAGGCTGGCGCTCAGTGTCCACA	4823
Qy	1058	AspSerGluValLeuAlaSerGluLysGlu-----LysAspValMetProLysGly	1074
Db	4824	GCCTCAGTGGCTACGAGCTCAITTCACAGAGCCACACAGATGATGTCTCCATCTCTG	4883
Qy	1075	ProValGlnGluAlaGlyAlaGluHis-----	1083
Db	4884	CATGCT---GAGGTGGCTCCCCACATCTCCACAGAAGTAGATGACTCCCTTTTCAGTGTCT	4940
Qy	1084	LeuAlaGlnGlySerGluThrGlyGlnAlaThr-----	1094
Db	4941	GTGTGTGAACACTACCACTACCATCTCAGGAACACAGAAATGCTCCATCTAAGAAGAATGC	5000
Qy	1095	-----	1104
Db	5001	CCAAGACCGATGTCAATTTCTCCACAGATTTCTCCCTAAACTGCAAAAGTCCAGGACA	5060
Qy	1105	AlaaspValAspHisValAlaThrCysGlnValLleLysLeu-----GlnGlnLeu	1121
Db	5061	CCCGTTCAAGATCACAGATCTGAACAGTCCTCAATGTCTATTGAATTTGGCCAGAATCT	5120
Qy	1122	MetGluGlnAlaValAlaProGluSerSerGluThrLeuThrAspSerGluThr-----	1139
Db	5121	CCTGAGCAATCCCTGCTATGACTTCAGTCGACAGTCTCCAGATCACCTACAGTGGGT	5180
Qy	1140	-----	1149
Db	5181	GCAGGCGTGTTCACATCACTGAAATGGGCCAATGAAAGTGGACTACAGTCTCTCTGAC	5240
Qy	1150	ThrAlaaspGlyThr-----GlnGlnAspGluThrLleAspSerGln	1163
Db	5241	ATGCAGGACTCCAGCTTTATCATCAATAGATACCACTATGGAGGAGCGCTCTACACCCAA	5300
Qy	1164	AspSerLysAlaThrAlaAlaValArgGlnSerGlnValThrGluGluGluAlaAlaThr	1183
Db	5301	GATAATGATCTTCTGAGCTCATCTCAGTATCTCAGGTA---GAGGCGTCCCGTCCACC	5357
Qy	1184	AlaGlnLysGluGluProSerThrLeuProAsnValProAlaGlnGluGluHisGly	1203

```

Db 5358 TCTTGTGCTATACCCCTTCTCAGATGCTCT-----CCTCTCCA----- 5399
QY 1204 GtugluProgluArgAspValleuGluProthrgluGluLeuThralaVal 1223
Db 5400 GAAAGTACTCTATCCGATGTGCTCTCCAGAGATATGCTTATATGCTCCTACAC 5459
QY 1224 ProvalleuAlaSerThrgluValGluGluGluValAspThreuleuSpglu 1243
Db 5460 -----TCTGAAAAAGTGCMAAGCTCGAAGAGAG 5489
QY 1244 LysValleuGluGluGluValPheValHisSerGlyProAsnSerGluValaVal 1263
Db 5490 AGCTCTCTCCAAATCTATATCTCTCCATCCACCCACGAGAGTCTCTCTTATAT 5549
QY 1264 AspValThrTyrAspSerGluValMetGlyVal-----AlaGlyCysGluGlu 1279
Db 5550 TCACCTACTTTTCAATCTACCTCTGCATCAAGAGAAAGAAACCACTTGCCACAG 5609
QY 1280 LysGluSerThrgluValGluSerLeuSerLeuGluGluGluMetGluThraSpVal 1299
Db 5610 TCCTCTTCCCAACATAGATGATCCGAGAGCCCTTATGCTTCGTCCTCAGTG 5669
QY 1299 ----- 1299
Db 5670 TTATTGCATACATGCAACACCTTACGCTTGAAATAGAGATTTCACACCTGCGCTG 5729
QY 1300 GluLysGluLysArgGluThrlLysProGluGluValSer-----GluGluGluGlu 1316
Db 5730 GAAAGAGACAGTGAGAGAGAACACCTGGTATGCTTATGCTATCAAAAGCTGAG 5789
QY 1317 GlnGluThralaAlaProGluHisGluGly-----ThrlGlyLysProVal 1332
Db 5790 GAAACCAACGAGTCCCAATGAAAGATGATGACTATGAGCTTATGAGAGACACC 5849
QY 1333 LeuThrlLeuAspMet-----ProSer 1339
Db 5850 CGAGCTTCAGATGTGGGTGCTTATGAGAAATAGAGAAACCAAAATCTCCA 5909
QY 1340 SerGluArgGlyLysAlaLeuGlySerLeuGlyGlySerProSerLeuProAspGlnAsp 1359
Db 5910 AGTGACAGTGGCTACTCTCATGAGACCATGGGAAACCTACCAAGACCTGAAAGTGT 5969
QY 1360 LysAlaGlyCysIleGluValGlnValGlnSerLeuAspThrlValThrglnThrala 1379
Db 5970 -----GACTATTCCTAT 5981
QY 1380 GluAlaValGluLysValIleGluThral-----ValIle 1391
Db 5982 GAAATTTATGAGAGAGACACACGAGCCCTGAAAGAGGCGGTACTCATATGACATAGT 6041
QY 1392 SerGluThrlGlyLysSerProGluCysValGlyAlaHisLeuLeuProAlaGluLysSer 1411
Db 6042 GAAAGAACCAACGAGCCCGGAGAGTGTACAGCTATGAAAAAGACTGAGAGGTCT 6101
QY 1412 -----SerAlaThrlGlyHisTrp 1418
Db 6102 AGAAGCTTCTGATGATACATCAGCAATGGCTATGATGACTGTGAGATGTGGCCAC-- 6158
QY 1419 ThrLeu-----GlnHisAlaGluAspThrlValProleuGlyProGlu 1432
Db 6159 ACACCTTGGGAGCCCACTGCTTATGAAACCACTGAGAAATTCACAGTTTCCCTGAG 6218
QY 1433 SerGlnAlaGluSerIleProIleIle----- 1441
Db 6219 TCTGAAGGTATTCATATGACATCAACAAGACAAACGAAACCCCTATCTTCACACA 6278
QY 1442 -----ValThrProAlaProGluSerThrlLeuHisProAsp 1453
Db 6279 TACTGTACAGAGCTCAGAGAAATACATAAGACCTCAGGATCCATCATATTCCTAC 6338
QY 1454 LeuGlnGlyGluLysSerAlaSerGlnArgGluArgSerGluGluLysLeuAsp 1473
Db 6339 GAGACTTCAGACTATGCTACTGCAAGAAAGAGTCCCTCAGAGAGCCGTCAGAGT 6398

```

```

QY 1474 AlaGly-----ProAspAlaAspGlyLysGlu 1482
Db 6399 GTGATTTATGCTCTGCTGCTCTTGTGAATACAGACACCCCAAGACAGACTTTCACC 6458
QY 1483 Ser-ThrlAlaIleGluLysValLeu----- 1490
Db 6459 TCTTTCATTAATCCCATTCCTTGTGAGTGTTCGCAATGAGAAACCCATGAAATCT 6518
QY 1491 -----LysAlaGluProGluIleLeuGluLeuGluSerLys 1504
Db 6519 GAAAGCCCTCCTCATCATCAAGAGAGAGCCCAAGCTCCAGAGAGAAACCAACAGGCG 6578
QY 1504 nLysIleValLeuAsnValIleGlnThralaValAspGlnPheAlaGtrGluThr-- 1523
Db 6579 CGACAGTGTGAAGAACCCCTCC--ACCTCAGTACAGCAGAGTCCATCCACAGCCGA 6637
QY 1524 -----AlaProGluThrlHisAlaTyrAspSerGlnThrgluValProAlaCysAr 1540
Db 6638 CTCTGATGTTCCTCCGAGAGACTGAAAGAGTCCCTTCATCAG-----GCCGATGCCA 6691
QY 1540 GlnAspSerArgGluProAsnArgCysTrpThrlLysMetLysAspAlaLysMetLysHis 1560
Db 6692 TATGACTGTGAAGACGAGTGGAAACCATCCCAAGACAAACTGTCAGTACAAACA 6751
QY 1560 S-----ProVal-----ProGlnProAlaGlu----- 1567
Db 6752 CATGACCCACTCCAGCTCCCGTGCMAAGACCGAGCCCTTCGACACCCCTGATGT 6811
QY 1568 -----AspLeuGlnValLeuThrlValLeuGlu-AlaTrpAlaGln----- 1580
Db 6812 GTCCATGTGTGACCCAGAGGCTTGCCATTGAGCAGAACTGGGCAAAAGCTTAAAGA 6871
QY 1581 -----ProArgLysCysLeuProArgLeuGlnLeuLysAlaPro 1593
Db 6872 AGATCTGAAGAGAGAACCAAAACCAAAAGCCAGGTACAAAGACCA 6918

```

Search completed: December 13, 2002, 00:45:32
Job time : 1027 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

QM protein - nucleic search, using frame_plus_p2n model

Run on: December 13, 2002, 00:24:46 ; Search time 124 Seconds
(without alignments)
3947.227 Million cell updates/sec

Title: US-09-902-432-4
Perfect score: 8073
Sequence: 1 MGAGSTGSRPEQAGSDT.....AWAQRKCLPRLQLKAPVSK 1596

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool/US0902432/runat 12122002 141452 25831/app query.fasta 1.1735
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0902432 @CGN 1 1 66 @runat 12122002 141452 25831 -NCPU=6 -ICPU=3
-NO XLPXY -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -PGAPOP=6 -PGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:.*
1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6005	74.4	5134	2	US-08-635-121-1
2	4281	53.0	6605	1	US-08-769-309A-4
3	4281	53.0	6605	3	US-08-994-570-4
4	462	5.7	5361	4	US-08-973-462-2
5	462	5.7	6152	4	US-08-973-462-1
6	421.5	5.2	6744	1	US-08-119-125A-2
7	415	5.1	6306	1	US-08-466-390-3
8	415	5.1	6306	1	US-08-470-950-3
9	415	5.1	6306	1	US-08-467-781-3
10	415	5.1	6306	2	US-08-483-924-3
11	412	5.1	6306	2	US-08-495-487-3
12	412	5.1	6306	5	PCT-US93-06160-3

13	410.5	5.1	6755	3	US-08-931-999-4	Sequence 4, Appli
14	388	4.8	11236	1	US-07-853-913-1	Sequence 1, Appli
15	366	4.5	9551	1	US-08-056-200-93	Sequence 93, Appli
16	366	4.5	9551	2	US-08-800-644-93	Sequence 93, Appli
17	365.5	4.5	6414	4	US-09-134-001C-1626	Sequence 1626, Ap
18	365.5	4.5	11091	4	US-09-134-001C-2243	Sequence 2243, Ap
19	357.5	4.4	7766	4	US-09-125-619-3	Sequence 3, Appli
20	354	4.4	8224	6	5180808-1	Patent No. 5180808
21	351	4.3	5661	4	US-08-938-105-2	Sequence 2, Appli
22	346.5	4.3	32207	2	US-08-770-379-20	Sequence 20, Appli
23	346.5	4.3	32207	4	US-08-757-669A-20	Sequence 20, Appli
24	346.5	4.3	32207	4	US-09-230-371A-20	Sequence 20, Appli
25	343.5	4.3	3337	1	US-08-072-610-1	Sequence 1, Appli
26	343.5	4.3	3337	2	US-08-719-822B-1	Sequence 1, Appli
27	343.5	4.3	3337	4	US-09-092-458-1	Sequence 1, Appli
28	339.5	4.2	3561	4	US-09-134-001C-1685	Sequence 1685, Ap
29	334.5	4.1	3489	2	US-08-728-323A-1	Sequence 1, Appli
30	334.5	4.1	3489	4	US-09-298-568-1	Sequence 1, Appli
31	333.5	4.1	4852	1	US-07-853-913-3	Sequence 3, Appli
32	321	4.0	5656	1	US-08-425-061-1	Sequence 1, Appli
33	321	4.0	5656	1	US-08-825-886-1	Sequence 3, Appli
34	321	4.0	5689	1	US-08-425-061-3	Sequence 3, Appli
35	321	4.0	5689	1	US-08-825-886-3	Sequence 3, Appli
36	321	4.0	5709	1	US-08-425-061-2	Sequence 2, Appli
37	321	4.0	5709	1	US-08-825-886-2	Sequence 2, Appli
38	321	4.0	5711	1	US-08-425-061-4	Sequence 4, Appli
39	321	4.0	5711	1	US-08-825-886-4	Sequence 4, Appli
40	321	4.0	5711	2	US-08-658-322-1	Sequence 1, Appli
41	321	4.0	5712	1	US-08-425-061-12	Sequence 12, Appli
42	321	4.0	5712	1	US-08-825-886-12	Sequence 12, Appli
43	321	4.0	5712	2	US-08-603-753D-1	Sequence 1, Appli
44	321	4.0	5712	3	US-09-009-753-1	Sequence 1, Appli
45	321	4.0	5712	4	US-08-986-106-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-635-121-1
; Sequence 1, Application US/08635121
; Patent No. 5910442
; GENERAL INFORMATION:
; APPLICANT: Gelman, Irwin H.
; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10112-0228
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/635,121
; FILING DATE: 19-APRIL-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Richard S
; REGISTRATION NUMBER: 26,154
; REFERENCE/DOCKET NUMBER: A30558 - 165/33603
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-408-2558
; TELEFAX: 212-765-2519
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
 LENGTH: 5134 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 FRAGMENT TYPE:
 ORIGINAL SOURCE:
 US-08-635-121-1

Alignment Scores:
 Pred. No.: 0 Length: 5134
 Score: 6005.00 Matches: 1232
 Percent Similarity: 97.33% Conservative: 7
 Best Local Similarity: 96.78% Mismatches: 27
 Query Match: 74.38% Indels: 14
 DB: 2 Gaps: 4

US-09-902-432-4 (1-1596) x. US-08-635-121-1 (1-5134)

QY 329 GtUyTthGluProAlaSerGluGluGluProAlaGluAspThrAspGlnAlaArg 348
 Db 2 GAAAGACAGAGCCAGCTCGAGAGAGCAGAGCCGGCAGAAACACACAGAGCCAG 61
 QY 349 LeuSerAlaAspTyrGluValGluLeuProLeuGluAspGlnValGluAspLeuGlu 368
 Db 62 TTGTACGACGACTACAGAGAGGTGAGCTGCTTGGAAAGACAGGTGTGTGCTGGAG 121
 QY 369 AlaSerSerGluGluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 388
 Db 122 GGATCCTCGAGAGAGAGGTGCTCTTGGCAACGAAAGTGTATAGAAAGATGGA 181
 QY 389 AlaHseGluValValAlaGluValHseValSerThrValGluValSerThrGluGlu 408
 Db 182 GCCCACCAGAAAGTTTGTCCAGAGGTCCAGCTGACACCTGTGAGAAAGACAGAGAG 241
 QY 409 GlnGlyGlyGlyGlyGluValGluGluGlyGlyValValGluGlyThrGlyGluSerLeu 428
 Db 242 CAGGAGAGAGAGAGAGCTGAAAGGGGCGTGTGTAGAAAGAACAGAGAAATCTTG 301
 QY 429 ProProGluLeuLeuAlaGluProGluGluValProGluGluAlaGluProAlaGlu 448
 Db 302 CCCCTCGAAGAACTGGCTAGCCCGAGAGGTCCCGAGAGCTGAGCTGTGAGAG 361
 QY 449 LeuMetLysSerArgGluMetCysValSerGlyGluAspHisThrGlnLeuThrAspLeu 468
 Db 362 CTGATGAAAGACAGAGAGATGTGTCTCTGAGAGAGACCACTGAACAGAGCTTA 421
 QY 469 SerProGluGluLysThrLeuProLysHisProGluGlyTLeuValSerGluValGluMet 488
 Db 422 AGCTCCGAAAGAGACGCTGCCCAACACCCAGAGGCAATGTCTGAGAGGTGAGATG 481
 QY 489 LeuSerSerGluGluArgLysValGlnGlySerProLeuLysLysLeuPheSerSer 508
 Db 482 CTGTCTCTCAGAGAAAGATCAAGGTACAGGGAGAGTCCCTTGAAGAACTCTTCACTAAC 541
 QY 509 SerGlyLeuLysLysLeuSerGlyLysLysGlnLysGlyLysArgGlyGlyGlyAsp 528
 Db 542 TCAAGGCTTAAAGAGCTGTCTGGAGAGAGCAGAGAGGAGAGAGAGGTGGGGAGAGAC 601
 QY 529 GtUgluProGluGluTyrGlnHisIleHisThrGluSerProGluSerAlaAspGluGln 548
 Db 602 GAAAGACCTCGAGAAATACCAACATTCACACCGAATCCCGCAGAGAGTGTATGAGCAG 661
 QY 549 LysGlyGluSerSerAlaSerSerProGluGluProGluGluThrCysLeuGluLys 568
 Db 662 AAGGGAGAGAGCTGTCTGTCTCCCGAGAGAGCTGAGAGAGACCACTGTCTGGAGAAA 721
 QY 569 GlyProLeuGluAlaProGluAspGlyGluAlaGluGluGlyThrThrSerAspGlyGlu 588
 Db 722 GGGCGCTGTGAGAGACCCAG-GATGGGAGAGCTGAGAGAGAACTTCTTCTGT--GGAGAG 778

QY 589 LysLysArgGluGlyTLeuThrProTPrAlaSerPheLysLysMetValThrProLysLys 608
 Db 779 AAGAGAGAGAGG--ATCATTCTCCGTGGGCATCTTCAAAAAAGTGTATACCCCAAGAAA 836
 QY 609 ArgValArgArgProSerGluSerAspLysGluGluGluGluValLysSerAla 628
 Db 837 CGGTC-CGAAAGCTTCTTGAGAGTACAGAGAGAGAGAGCTGAGAGAGTCAAGAGCC 895
 QY 629 ThrLeuSerSerThrAspSerThrValSerGluMetGlnAspGluValLysThrValGly 648
 Db 896 ACCTGTCTCTCACTGATAGCAGAGTGCAGAAATGCAAGATGAAGTCAAAACTTGTGT 955
 QY 649 GtUgluGlnLysProGluGluProLysArgArgValAspThrSerValSerTyrGluAla 668
 Db 956 GAGGAAACAAAGCCAGAGAAACAAAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1015
 QY 669 LeuLecCysValGlySerSerLysArgAlaArgLysValAspThrSerSerAspAspGlu 688
 Db 1016 CTGATTTGTGTGATCATTCAG 1075
 QY 689 GlyGlyProArgThrLeuGly-GlyAspSerHisArgAlaGluGluAlaSerLysAspLys 708
 Db 1076 ---GGGCCAAGGACCTGGAGGGGGACATGCACAGACAGAGAGAGGCGCAGAAAGACA 1132
 QY 708 GtUgluGlyThrAspAlaValProAlaSerThrGlnGluGluAspGlnAlaGlnGlySe 728
 Db 1133 AGAACCCCA-ACAGAGCTGTGTCTGCAGACCCAGAGAGAGAGAGAGAGAGAGAGAGAG 1191
 QY 728 rSerSerProGluProAlaGlySerProSerGluGlyGluGlyValSerThrTyrGluSe 748
 Db 1192 TTCTCACCAGCCAG 1251
 QY 748 rPheLysArgLeuValThrProArgLysLysSerLysSerLysLeuGluGluVal-- 767
 Db 1252 ATTTAAAGATTAGTCACTCCAGAAATAATCCAAATCGAAATCGAAAGAAAGAA--AG 1310
 QY 768 ----GluAspSerSerVal---GluGlnLeuSerThrGluTLeuGluProSerArgGlu 785
 Db 1311 CCGAAGAGACTTAATTTAG 1370
 QY 785 userTyrValSerLysLysPheLeuProGlyArgArgLysLysValAspGlyLys 805
 Db 1371 ATCTTGGGTTTCAATTAAAGAAATTCATCCCGAGAGAGAGAGAGAGAGAGAGAGAGAG 1430
 QY 805 s-GlnGluGlnAlaThrValGluAspSerGlyProValGluTLeuAsnGluAspAspPro 825
 Db 1431 GGCAGAAACAGCCACTGTGGAAGACTCAGGGCCAGTGGAGTAAATGAGAGCAGAGCTG 1490
 QY 825 snValProAlaValValProLeuSerGluTyrPheAlaValAlaGluArgGluLysMetGlu 845
 Db 1491 ATGTCCAGAGCGCTCGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1550
 QY 845 IagGlnGlyAspThrGluLeuProGluLeuLeuGluValValTyrValSerGluGluLys 865
 Db 1551 CCCAGGGAGATCGGAGCTGCCAGCTGCTGGGCTGTGTA---GTGTCCAGAGAGCTCA 1607
 QY 865 eLysThrLeuValHisThrValSerValAlaValIleAspGlyThrArgAlaValThr 885
 Db 1608 GTAAAGACTGTGTCACACCTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 1667
 QY 885 eValGluGluLysSerProSerTyrPheSerAlaSerValThrGluProLeuGluHisT 905
 Db 1668 GTGTGGAAGAGGCTGTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1727
 QY 905 hAlaGlyGluAlaMetProProValGluGluValThrGluLysAspLysIleIleAlaGlu 925
 Db 1728 CAGCGGAGAGAGCCATGTGCACTGTGAGAGGTCACTGAAAGAACATCATTTGAGAGAG 1787
 QY 925 LuThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHisAspAspMetValT 945
 Db 1788 AAACCTCTGTGTCTACCCAGAGAGTTCACAGAGAGGTAAAGATGCCATATACGACATGTCTA 1847

QY 945 hrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeuA 965
| | | | |
Db 1848 CCAGTGAAGTGGATTTCACCTCAGAGCTGTGACAGCCACAGAGACCTCAGAGCTCTCC 1907
| | | | |
QY 965 rgThrGluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSerAlav 985
| | | | |
Db 1908 GTACTGAAGAAGTTACCGAAGCATCGGGGCCGAAGAGACCACAGACATGCTGTCGCAG 1967
| | | | |
QY 985 alSerGlnLeuThrAspSerProaspThrThrGluGluAlaThrProValGlnGluValG 1005
| | | | |
Db 1968 TTTCCAGAGCTGACTGACTCCCCAGACACACAGAGGAGCCACCCAGTTCCAGAGGTAG 2027
| | | | |
QY 1005 luSerGlyValLeuAspThrGluGluGluGluArgGlnThrGlnAlaIleLeuGlnAlav 1025
| | | | |
Db 2028 AGGTGTGTGTGTAGATACAGAGAAGAGGAGGCCAGACGAGGCCATCTCTCAAGCCG 2087
| | | | |
QY 1025 alAlaAspLysValLysGluGluSerGlnValProAlaThrGlnThrValGlnArgThrG 1045
| | | | |
Db 2088 TTGCAGACAAGGTGAAGAGAGGTCCAGGTGCTGCAACCCAGACGTGTGCAGAGAACGG 2147
| | | | |
QY 1045 lySerLysAlaLeuGluLysValGluGluValGluGluAspSerGluValLeuAlaSerG 1065
| | | | |
Db 2148 GGTCAAAAGCACTGGAGAAGGTTGAGAGGTAGAGGAGTCCGAAGTGTGCTTCGG 2207
| | | | |
QY 1065 lylLysGluLysAspValMetProLysGlyProValGlnGluAlaGlyAlaGluHisLeuA 1085
| | | | |
Db 2208 AGAAAGAGAAGGAGCGTTATGCCGAAAGGACCCGTCGAGGAGCTGGAGCATCTTG 2267
| | | | |
QY 1085 laGlnGlySerGluThrGlyGlnAlaThrProGluSerLeuGluValProGluValThrA 1105
| | | | |
Db 2268 CACAGGGCTCTGAGACTGGACAGGCTACTCCAGAGAGCTTGAAGTTCCTGAAAGTCACAG 2327
| | | | |
QY 1105 laAspValAspHisValAlaThrCysGlnValLleLysLeuGlnGlnLeuMetGluGlnA 1125
| | | | |
Db 2328 CAGATGTAGACCATGTCCGACAGTGCAGGTATCAAGCTCCAGCAGTGTGGAAGCACTCCCT 2387
| | | | |
QY 1125 laValAlaProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrProL 1145
| | | | |
Db 2388 CCGTGGCCCTTGAGTCTATCCGAACCTTGACAGACAGTGAACAATGGAAGCACTCCCT 2447
| | | | |
QY 1145 euAlaAspSerAspThrAlaAspGlyThrGlnGlnAspGluThrIleAspSerGlnAsps 1165
| | | | |
Db 2448 TAGCAGATTACAGACTGCGAGTGGGACACAGCAGATGAACCATTCACAGCCAGGACA 2507
| | | | |
QY 1165 erLysAlaThrAlaAlaValArgGlnSerGlnValThrGluGluGluAlaAlaThrAlag 1185
| | | | |
Db 2508 GTAAAGCCACTGCAGCTGTCCAGGCAGTCACAGGTTCACAGAAGAAGAGCGCTACTGCTC 2567
| | | | |
QY 1185 lnlLysGluGluProSerThrLeuProAsnAsnValProAlaGlnGluGluHisGlyGluG 1205
| | | | |
Db 2568 AGAAAGAGGAGCCCTTCGACACTACCTAATATGTTCCAGCCCAAGGAACATGGGGAAG 2627
| | | | |
QY 1205 luProGlyArgAspValLeuGluProThrGlnGlnGluLeuThrAlaAlaAlaValProv 1225
| | | | |
Db 2628 AACAGGAAGAGATGTTCTTGAACCTACACAGCAGAGCTTGTCTGACGCCGTGCCCG 2687
| | | | |
QY 1225 alLeuAlaLysThrGluValGlyGlnGluGlyGluValAspTrpLeuAspGlyGluLysV 1245
| | | | |
Db 2688 TCTGGCAAAAGACTGAGTGGGTCAAGAGGTTGAGTGTGCTGGATGGAGAAAAG 2747
| | | | |
QY 1245 alLysGluGluGlnGluValPheValHisSerGlyProAsnSerGlnLysAlaAlaAspV 1265
| | | | |
Db 2748 TCAAAGAAAGAACAGGAGGTGTTGTATCACTCTGGACCCAAACAGTCGAAAGGCTGCTGATG 2807
| | | | |
QY 1265 alThrTrpAspSerGluValMetGlyValAlaGlyCysGlnGluLysGluSerThrGluV 1285
| | | | |
Db 2808 TGACATATGACGTGAAGTGAATGGAGTGGCCGGTGTACGAAAAGAGAGTACTGAAG 2867
| | | | |
QY 1285 alGlnSerLeuSerLeuGluGluGlyGluMetGluThrAspValGluLysGluLysArgG 1305
| | | | |
Db 2868 TGCAGAGCTCTAGCTGGAGGAGGAGATGGAACTGACGTGTGAAAGAGGAAAAGG 2927
| | | | |
QY 1305 luThrLysProGluGlnValSerGluGluGluGlnGluThrAlaAlaProGluHisG 1325
| | | | |

Db 2928 AGCAAAAGCCAGAGCAAGTGAAGAAGGTGAGCAAGAAACAGCCGCTCTCTGAGCATG 2987
| | | | |
QY 1325 luGlyThrTyrgLysProValLeuThrLeuAspMetProSerSerGluAtgGlyLysA 1345
| | | | |
Db 2988 AAAGGAACATACGGGAAGCCAGTCTTGACACTTGACATGCCAGCTCAGAGAGGGGAAGG 3047
| | | | |
QY 1345 laLeuGlySerLeuGlyGlySerProSerLeuProAspGlnAspLysAlaGlyCysIleG 1365
| | | | |
Db 3048 CACTGGGAAGCCTTGGAGGAAGCCCTTCTCCAGACCAAGACAAAGCAGGTTGCATAG 3107
| | | | |
QY 1365 luValGlnValGlnSerLeuAspThrThrValThrGlnThrAlaGluAlaValGluLysV 1385
| | | | |
Db 3108 AGGTTCAAGTTCAAGCTCGGACACACAGTCACTCAACAGCAGAGCTGTGGAAAGG 3167
| | | | |
QY 1385 alIleGluThrValValIleSerGluThrGlyGluSerProGluCysValGlyAlaHisL 1405
| | | | |
Db 3168 TCATAGAAGCGTGTGATTTCCAGACAGAGTGAAGTCCAGAGTGTGTAGGTGCACACT 3227
| | | | |
QY 1405 euLeuProAlaGluLysSerSerAlaThrGlyGlyHisTrpThrLeuGlnHisAlaGluA 1425
| | | | |
Db 3228 TATTACAGCTCAGAAAGTCTCTGCAACGGGTGGCCACTGGACTCTTCAGCATGCAGAG 3287
| | | | |
QY 1425 spThrValProLeuGlyProGluSerGlnAlaGluSerIleProIleIleValThrProA 1445
| | | | |
Db 3288 ACACGGTACCCCTGGGGCCCTGAGTCTCAGGCAGAAATCCATCCCAATCATAGTAACCTCTG 3347
| | | | |
QY 1445 laProGluSerThrLeuHisProAspLeuGlnGlyGluIleSerAlaSerGlnArgGluA 1465
| | | | |
Db 3348 CTCTGAAAGCACCTTACATCTCTGACCTACAGGAGAGAAATAGCGCATCCAGAGAGAGC 3407
| | | | |
QY 1465 rGserGluGluGluAspLysProAspAlaGlyProAspAlaAspGlyLysGluSerThrA 1485
| | | | |
Db 3408 GATCAGAGGAAGAGGACAAAGCCAGATGCTGCTCTGATCTGACGGCAAGAGAGTACAG 3467
| | | | |
QY 1485 laIleGluLysValLeuLysAlaGluProGluIleLeuGluLeuGluSerLysSerAsnL 1505
| | | | |
Db 3468 CAATCAGCAAAAGTCTCAAGGCTGAAACCTTGAGATCTCTGAACTTGAGAGTAAGAGCAACA 3527
| | | | |
QY 1505 ysIleValLeuAsnValIleGlnThrAlaValAspGlnPheAlaArgThrGluThrAlaP 1525
| | | | |
Db 3528 AGATTGTCTGAACGTCATTTCAGACAGCCGTTGACCAAGTTGCGACCTGACAGAAACAGCCC 3587
| | | | |
QY 1525 roGluThrHisAlaLysAspSerGlnThrGlnValProAlaCysArgLeuAspSerArgG 1545
| | | | |
Db 3588 CCGAAACTCATGCTTATGATTTCACAGACCCAGGTTCTGCAATGCGCTTGACAGCAGGG 3647
| | | | |
QY 1545 luProAsnArgCysTrpThrLysMetLysAspAlaLysMetLysHisProValProGlnP 1565
| | | | |
Db 3648 AGCCCAACAGATGCTGACCAAAATGAAAGTTGCCAAGATGAAACACCCAGTCCCGCAGC 3707
| | | | |
QY 1565 roArgGluAspLeuGlnValLeuThrValLeuGluAlaTrpAlaGlnProArgLysCysL 1585
| | | | |
Db 3708 CCAGAGAGACTTGCAGTCTCTGACCGTCTTGAGGAGCATGGC--TCAGCTCGGAAATGCT 3765
| | | | |
QY 1585 euProArgLeuGlnLysAlaProValSerLys 1596
| | | | |
Db 3766 TGCCGCGCTTGCAAGTTGAAAGCGCCGCTGTCAAAG 3800
| | | | |
RESULT 2
US-08-769-309A-4
; Sequence 4, Application US/08769309A
; Patent No. 5741890
; GENERAL INFORMATION:
; APPLICANT: Scott, John D.,
; APPLICANT: Nauer, Brian J.,
; APPLICANT: Klauck, Theresa M.
; TITLE OF INVENTION: Protein Binding Domains of Gravin
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower/233 South Wacker Drive
; CITY: Chicago

```

STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/769,309A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 5741890and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 27866/33451
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 6605 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 192..5531
US-08-769-309A-4

Alignment Scores:
Pred. No.: 1,51e-288 Length: 6605
Score: 4281.00 Matches: 965
Percent Similarity: 68.69% Conservative: 189
Best Local Similarity: 57.44% Mismatches: 409
Query Match: 53.03% Indels: 117
DB: 1 Gaps: 37

US-09-902-432-4 (1-1596) x US-08-769-309A-4 (1-6605)
QY 1 MetGlyAlaGlySerThrGlnArgSerProGluGln---ProAlaGlySerAsp 19
DB 192 ATGGGGCGCGGGAGCTCCACCGAGCGCAGCGCCGAGCGCGCCGAGGGAGACTCC 251
QY 20 ThrProSerGluLeuValLeuSerGlyHisGlyProAlaAlaGluLaser---GlyAla 38
DB 252 AGCCCGGCTGAGCCCGAGCCCGAGCGCGCGCCCTCGGCGGAGCGCGCCAGACCC 311
QY 39 AlaGlyAspProAlaAspAla-----AspProAlaThrIleuProGlnIlyAsnGly 56
DB 312 ACCGGGACCCCGCCATCGCTGCTCGGACCCGCGACCCGACCAAGCTCTACAGAAATGGT 371
QY 57 GlnLeuSerValAsnGlyValAlaGluGlnGlyAspValHisValGlnGluGluAsn 76
DB 372 CAGCTGTCCACATCAATGAGCTGAGCAGATGAGTCAAGCTCCAGGAGGGTGC 431
QY 77 GlnGluGlyGln-----GlnGlu 82
DB 432 CTAAATGGCCAGAAAGAGAGCCCTGAACGGTCAAGAGCCCTAAACAGCCAGAGAAAGAA 491
QY 83 GluValAlaAspGluAspValGlyGlnArgIlyuSerGluAspValArgGluIlyAspArg 102
DB 492 GAAGTATTGTACGAGAGGTGGACAGAGACTCTGAAGATGTGAGCGAAAGACTCC 551
QY 103 ValGluGluMetAlaAlaAsnSerThrAlaValGluAspIleThrIlyAspGlyGlnGlu 122
DB 552 GATTAAGAGATGGCTACTAAGTCAGCGGTTGTTCACAGACTCAACAATGATGGCGAG 611
QY 123 GlnThrSerGluIleIleGluGlnIleProAlaSerGluAsnAsnValGluGluMetVal 142
DB 612 GAGAACCGAAAT--ATCGAAGAGATTCCTTCTTCAGAAAGCAATTAGAGAGCTAAC 668

```

```

QY 143 GlnProAlaGluSerGlnAlaAsnAspValGlyPheIlySlyValPheIlyPheValGly 162
DB 669 CAACCCACTGAGTCCACAGCTTAATGATTTGAAGAGGTGTATTAAAGTTGTGTGGC 728
QY 163 PheIlyPheThrValIlyValAspIlyAsnGlyIlySerAspThrValGluLeuThr 182
DB 729 TTTAATTTCAGTGTAAAGATTAAGACAGAGAGCTGACACTGTCCAGCTACTACT 788
QY 183 ValIlySlyAspGluGlyGluGlyValAlaGluLaserValGlyAlaGlyAspHisGlnGlu 202
DB 789 GTGAAGAAAGATGAAGGGAGGAGCA-----GAGGGGGCTGGGACACACAGAGAC 839
QY 203 ProSerValGluThrAlaValGlyGluSerAlaSerIlySlySerGluLeuIlyGlnSer 222
DB 840 CCCACCTT-----GGGGCTGGAAGAACACATCCAAAGAAAGGAAAGCAACCAATCT 893
QY 223 ThrGluIlySlnGluGlyThrIleuIlySlnGluIlyGlnSerThrGluIleProLeuGln 242
DB 894 ACAGAGAAACCCGAGAGAGACCTGAAGCGTGGCAAGCCACGCAAGAAATTTCTCCCA 953
QY 243 AlaGluSerAspGlnAlaAlaGluGluGluAlaIlyAspGluGlyGluGluIlySlnGlu 262
DB 954 GCCGATCTGGCCAAAGCAAGTG--GAGGATGCAAAGAGAGAGAGAGAGAAACAGAA 1010
QY 263 LysGluProThrIlySerProGluSerProSerSerProValAsnSerGluThrThrSer 282
DB 1011 AAAGAACTTAGCAAGTCTGCAGAAATCTCCGACTACTCCCGTGCAGTGAACAGAGATCA 1070
QY 283 SerPheIlySlyPhePheThrHisGlyTrpAlaGlyTrpArgIlySlyThrSerPheIly 302
DB 1071 ACCTTCAAAAATTTCTTCACTCAAGTGGCGGCTGGCCGCAAAAACAGTTTTCAGG 1130
QY 303 IlySerIlySlyAspAspLeuGluThrAlaGluIlyArgIlySlyGlnGluAlaGluIly 322
DB 1131 AAGCGAAGAGAGAGAGAGTGAAGTGAAGCTTCAGAGAGAAAGAAAGAAAGCAAGCA 1190
QY 323 ValAspGluGluGluIlySlyGluIlyThrProAlaSerGluGlu----- 337
DB 1191 GTAGACACAGAAAGAGAGAGAGAGAGAGTGGCTCCGAGAAACAGCCCTCCGAG 1250
QY 338 -----GlnGluProAlaGluAspThrAspGlnAlaArgLeuSerAlaAspTrp 353
DB 1251 CAAGCCACCCACAGAGAGCGGACAGAAAGTGCACAGAGCCCGGTATTCAGCTAAAT 1310
QY 354 GlnIlyValGluLeuProLeuGluAspGlnValGlyAspLeuGluAlaSerSerGluGlu 373
DB 1311 GAGAAAGTTGAGCTGCCCTCAGAGAGAGCAAGTCAAGTGGCTGCAGAGACCTTCTGAAG 1370
QY 374 LysCysAlaProLeuAlaThrGluValPheAspGluIlyMetGluAlaHisGln---Glu 392
DB 1371 AAACCTGCTCCCTTGGCCACAGAAAGTGTGATGAGAAATAGAAAGTCCACCAAGAG 1430
QY 393 ValValAlaGluValHisValSerThrValGluIlySlnGluGluGlnGlyGly 412
DB 1431 GTTGTGGCCGAAGTCCAGTCAACCTGGAGAGAGAAACGAAAGAGAG----- 1481
QY 413 GlyIlyAlaGluGlyValValValAlaGluIlyThrGlyGluSerLeuProGluIly 432
DB 1482 -----AAACGAGGTGGAAGAAACAGCAGAGGTCTGTGCCAGCTGAGAA 1526
QY 433 LeuAlaGluProGlnGluValProGlnGluAlaGluProAlaGluGluLeuMetIlySer 452
DB 1527 TTGGTTGGATGATGACAGAACTCAGAGAGCGGAACCTGCACAGAGACTGTGTGAAGTC 1586
QY 453 ArgGluMetCysValSerGlyIlyAspHisThrGlnLeuThrAspLeuSerProGluGlu 472
DB 1587 AAGAGAAAGCTGTGTTCGGAAGAGACCTTACACAGGAGCTGACGTCTGATGAG 1646
QY 473 LysThrLeuProIlyHisIleProGluIlyIleValSerGluValGluMetLeuSerGln 492
DB 1647 AAGGTGCTGTCCAAACCCCGGAGAGCGGTGTGATGAGGTGGAATGTGTCAACAG 1706

```

QY	493	GluArgIleLysValGlnGlySerProLeuLysLysLeuPheSerSerSerGlyLeuLys	512
DB	1707	GAGAGAATGAAGGTGCAGGGNAAGTCACACTAAAGAAGACTTTTACCAGCACTCGGCTTAAAA	1766
QY	513	LysLeuSerGlyIlyValLysGlnLysGlyLysArgGlyGlyGlyLysAspGluGluProGly	532
DB	1767	ANGCTTTCTGAAAGAACAAGAANGGAAGAAGA---GGAGGAGGAGACGAGGAATCAGGG	1823
QY	533	GluTrpGlnHisIleHisThrGluSerProGluSerAlaAspGluGlnLysGlyLeuSer	552
DB	1824	GAGCACACTCAGGTTCACGCCGATCTCCGACAGCCAGGAGACCAAAAGGCGAGAGC	1883
QY	553	SerAlaSerSerProGluGluProGluGluThrThrCysLeuGluLysGlyProLeuGlu	572
DB	1884	TCTGCCCTCATCCCTGAGGAGCCGAGGAGATCACTGTCTCGAAAGAGGCTTACGCCGAG	1943
QY	573	AlaProGlnAspGlyGluAlaGluGluGlyThrThrSerAspGlyGluLysLysArgGlu	592
DB	1944	GTGCAGCAGATGGGGAAGCTGAAGAAGAGAGTACTTCCGATGGAGAGAAAAAAGAGAA	2003
QY	593	GlyIleThrProTrpAlaSerPheLysLysMetValThrProLysLysArgValArgArg	612
DB	2004	GGTGTCACTCCCTGGGCATCATTTCAAAGAAGTGTGACGCCCAAGACGGTGTAGACGG	2063
QY	613	ProSerGluSerAspLysGluGluLeuGluLysValLysSerAlaThrLeuSerSer	632
DB	2064	CCTTCGGAAGAGTGATAAAGAAGATGAGCTGGACCAAGGTCAAGAGCGCTACCTGTCTCTCC	2123
QY	633	ThrAspSerThrValSerGluMetGlnAspGluValLysThrValGlyGluGluGlnLys	652
DB	2124	ACCAGAGACACAGCCTCTGAAATGCAAGAAGAANAATGAAGGGAGCGCTGGAGAGCCAAAG	2183
QY	653	ProGluGluProLysArgValAspThrSerValSerTrpGluAlaLeuIleCysVal	672
DB	2184	CCGGAAGAACCAAGCCAGGTGGATCTCAGTATCTTTGGGAAGCTTTAATTTGTGTG	2243
QY	673	GlySerSerLysLysArgAlaArgLysAlaSerSerSerAspAspGluGlyGlyProArg	692
DB	2244	GGATCATCCAAGAAAAAGACGACGACGAGGTCTCTTCTGATGAGCAAGGGGGACCAAAA	2303
QY	693	ThrLeuGlyGlyAspSerHisArgAlaGluGluAlaSerLysAspLysGluAlaGlyThr	712
DB	2304	GCAATGGGAGGAGACCAACCAAGAGCTGATGAGCGCGGAAAAAGACAAAGACACGGGACAC	2363
QY	713	AspAlaValProAlaSerThrGlnGlnAspGlnAlaGlnGlySerSerSerProGlu	732
DB	2364	GACGGGATCTCTGCTGTTCCCAAGAACATGATCCAGGCGACGGAAGTCTCTCCCGGAG	2423
QY	733	ProAlaGlySerProSerGluGlyGluGlyValSerThrTrpGluSerPheLysArgLeu	752
DB	2424	CAAGCTGGAAGCCCTACCGAAGGGGAGGGCGTTTCCACCTGGGAGTCTATTTAAAGGTTA	2483
QY	753	ValThrProArgLysLysSerLysSerLysLeuGluGluLysAlaGluAsp-----	769
DB	2484	GTACGCCCAAGAAAAAATCAAAAGTCCAAGCTCGAAGAGAAAAAGCGAAGATCCTATAGCT	2543
QY	770	--SerSerValGlnLeuSerThrGluIleGluProSerArgGluGluSerTrpVal	788
DB	2544	GGGTCTGTTAGAACATTCCTCCACAGACTGAAACCCCGTGAAGAGAAATCTCTGGTCT	2603
QY	789	SerLysLysPheIleProGlyArgArgLysLysArgAlaAspGlyLysGlnGluGln	808
DB	2604	TCAATCAAGAAGTTTATTCCTGGACGAGGAAGAAAAAGCCAGATGGCAAAACAAGACAA	2663
QY	809	AlaThrValGluAspSerGlyProValGluIleAsnGluAspAspProAsnValProAla	828
DB	2664	GCCCCCTGTGAAGACGACGAGGGCCCAACAGGGGCCCAACGAAGATGATCTCTGATGTCCTCGGCC	2723
QY	829	ValValProLeuSerGluTrpAsnAlaValGluArgGluLysMetGlu-----Ala	845
DB	2724	GTGTCTCCTCTGTCTGAGTATGATCTGTAGAAAGGGAGAAATGGAGGCACAGCAAGCC	2783
QY	846	GlnGlyAsnThrGluLeuProGlnLeuLeuGlyAlaValTrpValSerGluGluLeuSer	865

Db	2784	CAAAAGGCGCAGCAGCAGCGCCGAGCAGGAAGGCAGCCACTGAGGTGTGTCCAAAGAGCCTCAGC	2843
Qy	866	LyethrLeuValHisThrValSerValAlaValIleAspGlyThrArgAlaValThrSer	885
Db	2844	GAGAGTCAGGTTCATATGATGGCAGCAGCTGTCGTGACGGACGAGGCGAGCTACCAT	2903
Qy	886	ValGluGluArgSerProSerTrpIleSerAlaSerValThrGluProLeuGluHisThr	905
Db	2904	ATTGAAGAAAGGTCTCTCTCTTGGATATCTGCTTCAGTGCAGAACTCTTTTGAACAAGTA	2963
Qy	906	AlaGlyGluAlaMetProProValGluGluValThrGluLysAspIleIleAla---Glu	924
Db	2964	GAAGCTGAAGCCGCACTGTTAACTCAGGAGGTATTGGAAAGAGAGAAGTAATTTGCAAGAA	3023
Qy	925	GluThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHisAspAspMetVal	944
Db	3024	GAACCCCCACGGTTACTGAACCTCTGCGCAGAGAACAGAGAGGCCCGGGGCGACACCGTC	3083
Qy	945	ThrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeu	966
Db	3084	GTTAGTGAGGCGGAATTGACCCCGGAAGCTGTGACAGCTGCAGAACTGCAGGGCCATTG	3143
Qy	965	ArgThrGluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSerAla	984
Db	3144	GGTTCCGAAGAGGAACCGAAGCATCTGCTGCTGAAGAGACACAGAAATGGTCTCAGCA	3203
Qy	985	ValSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGluVal	1004
Db	3204	GTCTCCCGATTAAACCGACTCCCGACACACACAGAGGAGGCCACTCCGGTCGAGAGGTG	3263
Qy	1005	GluSerGlyValLeuAspThrGluGluGluArgGlnThrGlnAlaIleLeuGlnAla	1024
Db	3264	GAAGTGGCGTACCTGCATAGNAGCAAGAGAGCGGACTCAAGAGGTCTCTCCAGGCA	3323
Qy	1025	ValAlaAspLysValLysGluLysGlnValSerGlnValProAlaThr-----	1038
Db	3324	GTGCGCAAAAAGTGAAGAGGAATCCCAAGTGCCTCGGCAGCGGTGGGCCAGAGAATGTG	3383
Qy	1039	---GlnThrValGlnArgThrGlySerLysAlaLeuGluLysValGluGluValGluGlu	1057
Db	3384	CTTCAGCCTGTGCAGAGACA-----GAGCGCAAGAAGACCAAGAGAG	3425
Qy	1058	AspSerGluValLeuAlaSerGluLysGluLysAspValMetProLysGlyProValGln	1077
Db	3426	CAGGCTGAAGCGTCGGGTCTGAAGAAAGAGCGGATGTAGTGTTCAAAGTAGATGCTCAG	3485
Qy	1078	GluAlaGlyAlaGluHisLeuAlaGlnGlySerGluThrGlyGlnAlaThrProGluSer	1097
Db	3486	GAGCGAAAAAACTGAGCTTTTACACAAGGGAAGGTGTGGGGCGACACCACCCCAAGAAAGC	3545
Qy	1098	LeuGlu---ValProGluValThrAlaAspValAspHis-----ValAlaThrCys	1113
Db	3546	TTTGMAAAAGCTCCTCAAGTCACAGAGACATAGTAGTCCAGTGAGCTTGTAAACCACTGTG	3605
Qy	1114	GlnVal-----11eLysLeuGlnGlnLeu---MetGluGlnAlaValAla	1127
Db	3606	CAAGCCGAACCTTAGCTGGGTAAANAATCACAGGAGATGTTGATGGAACAGGCTATCCCC	3665
Qy	1128	ProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrProLeuAlaAsp	1147
Db	3666	CCTGACTCGGTGGAACCCCTACAGACAGTGAGACTGATGGAAGCAGCACCCTCGTAGCCGAC	3725
Qy	1148	SerAspThrAlaAspGlyThrGlnGlnAspGluThrIleAspSerGlnAspSerLysAla	1167
Db	3726	TTTGACCCACCGGCACAACCCAGAAAGACGAGATTGGAAATCCATGAGGAGAATGCAG	3785
Qy	1168	ThrAlaAlaValArgGlnSerGlnValThrGluGluGluAlaAlaThrAlaGlnLysGlu	1187
Db	3786	GTGCATCTGGTA---CCAGTCAGGGGCACAGAGCAGAGGCGAGTTCTCTGCACAGAAAGAG	3842
Qy	1188	GluProSerThrLeuProAsnValProAlaGlnGluGluHisGlyGluGluProGly	1207

```

Db 3843 AGGCTCCAGCA---CCTTCCAGTTTGTGTTCCAGAGAAACTTAAGAACATCAAG 3899
Qy 1208 ---ArgspvalleugluProthrglngluLeuThralaalaProvalleu 1226
Db 3900 ATGGAAGACACTCTGAGCATACAGTAAAGAGTGCTAGTGAAGATCTATCCATTCTG 3959
Qy 1227 Alalystrhgluvalglngluvalaertrpleuaspelglulysvallys 1246
Db 3960 TCMAAGACTGAGGGGAGCTCA-----GAGGCTGACCACTATGCTGATGAAGAAACCAA 4013
Qy 1247 -----GluGluglngluValPheValHisSerGly-----Proanser 1259
Db 4014 GACGTAACATTTTTCAGAGAGCTTGAGGGGCTATAGACACAGGCTAACAGCTAGTCCG 4073
Qy 1260 GlulysalaalaAspValThrTyraSerSerGluValMetGlyValaIglYcseglngu 1279
Db 4074 GAAAGGTGACAGAGATGCCCTTAAGGTGAAGGACAGAGAGAGCTGAATGTAAAG 4133
Qy 1280 LysgluSerThrgluValGlnSer-----LeuSerleugluGluGlu 1294
Db 4134 GATGATGCTCTTGAACTGACAGAGCTGACGCTAAGTCTCTCCATCCCGCTGAGAGAGAG 4193
Qy 1295 MetGluThraAspValglulysArggluThraLysProgluGluValserGluGlu 1314
Db 4194 ATGTAGTTCATCAAGTGAAGAGAGAGAAACAGAGACAGACCAATGATGAAAGAG 4253
Qy 1315 Gly---GluGlngluThraIalaIProgluHisgluGlyThrglyLysProvalleu 1333
Db 4254 AAGCTTGACAGAGAAACAGCTGTACCTGATGTAAGAGAGTCAAGTCAAGAGCTCCAG 4313
Qy 1334 ThrLeuAspMetProSerSerSerGluValGlyValaLeuGlySerleuglySer--- 1352
Db 4314 ACAGTGCAATGTGCCCATCATAGATGGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4373
Qy 1353 ProSerLeuProAspGlnAspLysAlaGlyCysIleGluValGlnValGlnSerLeuAsp 1372
Db 4374 CCTCCCTGCCCTAGGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4433
Qy 1373 ThrThraValThrglnThraIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 1390
Db 4434 GGAATCATTCACCTCAACAGCGGCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4493
Qy 1391 HisSerGlnThrglyGluSerProgluGlyValaGlyAlaHisleuLeuProalagluLys 1410
Db 4494 ATTTTAGAAACAGGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4553
Qy 1411 SerSerAlaThrglyGlyHisTrpThraLeuGlnHisAlaGluAspThraIaIaIaIaIa 1430
Db 4554 TCCTCTGAAAAAATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4613
Qy 1431 ProgluSerGlnAlaGluSerIleProIleIleValThraProAlaProgluSerThraLeu 1450
Db 4614 CCGGACTGTCAGGCAAAATCGACACAGAGATGATGTCGTAATCGAAGAAAGAGCTTA 4673
Qy 1451 HisProAspLeuGlnGlyLysIleSerAlaSerGlnArgGluArgSerGluGluAsp 1470
Db 4674 ACTTCGAGCTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4733
Qy 1471 LysProAspAlaGlyProAspAlaAspGlyLysGluSerThraIaIaIaIaIaIaIaIa 1490
Db 4734 GAGCAGAGTTCCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4784
Qy 1491 LysAlaGluProglu-----IleLeuGluLeuGluSerLysSerAsnLysIleValleu 1508
Db 4785 -----GAGCCTGAAAAAATGGAGATTTTGAACTTGAGACCAAAAGAGAGAGAGAGAG 4838
Qy 1509 AsnValIleGlnThraIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 1527
Db 4839 AACATCATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4898
Qy 1528 HisAlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 1547
Db 4899 TTGACGCTGAGTTAGACAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4958

```

```

Qy 1548 ArgCysTrpThrLys-----MetLysAspAla-LysMetLysHisProVal 1562
Db 4959 CAGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5018
Qy 1562 LProgluProArgGlu-----AspLeuGlnValleuThraIaIaIaIaIaIaIaIaIa 1575
Db 5019 ACTTGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5078
Qy 1575 uGluAlaTrpAlaGlnProArgLysCysLeuProArgLeuGlnLeuLysAlaProVal 1594
Db 5079 AAAGCATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5136

RESULT 3
US-08-994-570-4
Sequence 4, Application US/08994570
Patent No. 6090929
GENERAL INFORMATION:
APPLICANT: Scott, John D.,
APPLICANT: Nauert, Brian J.,
APPLICANT: Klauk, Theresa M.,
TITLE OF INVENTION: Protein Binding Domains of Gravin
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower/233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/994,570
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: No. 6090929and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 27866/33451
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 6605 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 192..5531
US-08-994-570-4

Alignment Scores:
Pred. No.: 1.51e-288
Score: 4281.00
Percent Similarity: 68.6%
Best Local Similarity: 57.44%
Query Match: 53.03%
Gaps: 37
US-09-902-432-4 (1-1596) x US-08-994-570-4 (1-6605)
Qy 1 MetGlyAlaGlySerSerThrgluGlnArgSerProgluGln---ProAlaGlySerAsp 19
Db 192 ATGGGCGCCGGAGAGCTCCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 251

```


Db 2364 GACGGATCTTCTGCTGCTCCCAAGAACTGATCCAGGGCAGGAGTTCTCCCGGAG 2423
 Qy ProAlaGlySerProSerGluGluGluValSerThrTyrGluSerPheLeu 752
 Db 2424 CAAGCTGGAGGCTTACCGAAGGGAGGGGCTTCCACTGGGAGTCATTAAAGTTA 2483
 Qy ValThrProArgLysLysSerLysSerLysLeuGluGluValAlaGluAsp 769
 Db 2484 GTCAGCGCCAAAGAAAAAACAAGTCCAGCTGGAAGAAAAAGCAACATCCATAGCT 2543
 Qy ---SerSerValGluGluLeuSerThrGluLeuGluProSerArgGluGluSerTyrAl 788
 Db 2544 GGGTCTGGGTGAACATTCCTCCAGACTGAAACCCGGTAAAGAAATATCTGGGCT 2603
 Qy SerLysLysPheLeuProGluArgArgLysArgAlaAspGlyLysGluGluGln 808
 Db 2604 TCAATCAAGAAATTATCTCTGACGAAAGAAAGAAAGCCAGATGGGAAACAGAACAA 2663
 Qy AlaThrValGluAspSerGlyProValGluLeuAsnGluAspProAsnValProAla 828
 Db 2664 GCGCTGTGAAGACGACGAGGCCAACAGGGGCCAACGAAGATGACTCTGATGCTCCGGCC 2723
 Qy ValValProLeuSerGluTyrAsnAlaValGluArgLysMetGlu-----Ala 845
 Db 2724 GTGGTCCCTCTGTGAGTATGCTGTGAAAGGAGAAATGGAGGACACAGCAAGCC 2783
 Qy GlnGlyAsnThrGluLeuProGluLeuGluAlaValTyrValSerGluGluLeuSer 865
 Db 2784 CAAAAGGCGCAGAGACGCCGAGAGAGGAGGCCACCTGAGAGTGTCGAAGGCTCAAC 2843
 Qy LysThrLeuValHisThrValSerValAlaValLeuAspGlyThrArgAlaValThrSer 885
 Db 2844 GAGAGTCAGGTTCATATGATGACGACGATGTCGCTGAGAGGAGCGAGGCGAGCTACCAT 2903
 Qy ValGluGluArgSerProSerTyrPileserAlaSerValThrGluProGluGluHisThr 905
 Db 2904 ATTGAAGAAAGTCTCTCTTGTGATATCTGCTTCAGTACAGAACCTTTGAACAAGTA 2963
 Qy AlaGlyGluAlaMetProProValGluGluValThrGluLysAspLeuAla---Glu 924
 Db 2964 GAAGCTGAAGCCGCACTGTAACTGAGAGGATATTGAAAGAAAGTAATTGACAGAA 3023
 Qy GlnThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHisAspMetVal 944
 Db 3024 GAACCCCGCCAGGTACTCAACTCTGCCAGAAACAGAGAGGCCCGGGCCACAGCTC 3083
 Qy ThrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeu 964
 Db 3084 GTTAGTGAAGCGGAATTGACCCCGAAGCTGTACAGCTGCAGAAATCTGACGGGCCATTG 3143
 Qy ArgThrGluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSerAla 984
 Db 3144 GGTTCGAGAAAGAAACGCAAGCATCTGCTGAGAGAGCCACAGAAATGGGTGACAGA 3203
 Qy ValSerGluLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGluVal 1004
 Db 3204 GTCTCCAGATTAAACCACTCCCGACAGACACACAGAGAGGCCACTCCGGTGCAGGGGTG 3263
 Qy GlnSerGlyValLeuAspThrGluGluGluGluArgGlnThrGlnAlaAlaLeuGlnAla 1024
 Db 3264 GAAGGTGGGTACTGACATAGAAAGAGCAAGAGGGCGGACTCAAGAGGCTCTCCAGGCA 3323
 Qy ValAlaAspLysValLysGluSerGlnValProAlaThr----- 1038
 Db 3324 GTGGCAGAAAATGTAAGAAAGAAATCCAGCTCCCTGCGACCGGTGGCCAGAAATGG 3383
 Qy ---GlnThrValGlnArgThrGlySerLysAlaLeuGluLysValGluValGluGln 1057
 Db 3384 CTTCAAGCTGTGACAGAGCA-----GAGCGCAAGAAAGCCAGAGAG 3425
 Qy AspSerGluValLeuAlaSerGluLysGluLysAspValMetProLysGlyProValGln 1077
 Db 3426 CAGGCTGAAGCGTGGGTCTGAAGAAAGAGACGAGATGTAGTTGAAGATGATGCTCAG 3485

Qy GlnAlaGlyAlaGluHisLeuAlaGlnGlySerGluThrGlyGlnAlaThrProGluSer 1078
 Db 3486 GAGCGAAAACCTTATGAGCCCTTTTACACAAAGGAAGGTGGTGGGACAGACCCCGAAGAC 3545
 Qy LeuGlu---ValProGluValThrAlaAspValAspHis-----ValAlaThrCys 1113
 Db 3546 TTTGAAAAAGCTCTCTCAAGTACAGAGAGCATAGATGCCACTGACTTTGTAACCACTTGT 3605
 Qy GlnVal-----IleLysLeuGlnGluLeu---MetGluGlnAlaValAla 1127
 Db 3606 CAAGCCGAACCTTAGCTGAGGTAAATCAACAGAGATGTGATGAAACAGGCTATCCCC 3665
 Qy ProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrProLeuAlaAsp 1147
 Db 3666 CCTGACTGGGTGAACCCCTTACAGACAGTGAAGACTGATGGAACGACCCCTGTAGCCAC 3725
 Qy SerAspThrAlaAspGlyThrGlnGluAspGluThrIleAspSerGlnAspSerLysAla 1167
 Db 3726 TTTGACGACCAAGGACCAACCCAGAAAGACGATGTGGAAATCCATGAGAGAAATGAG 3785
 Qy ThrAlaAlaValArgGlnSerGlnValThrGluGluGluAlaAlaThrAlaGlnLysGlu 1187
 Db 3786 GTGCATCTGTA---CAAGTCAAGGGGACACAGAAAGAGGACAGTTCCGCAACAGAAAG 3842
 Qy GlnProSerThrLeuProAsnAsnValProAlaGlnGluGluHisGlyGluGluProGly 1207
 Db 3843 AGGCTCCAGCA---CCTTCCAGTTTGTGTTCAGAGAAAGAACTAAAGAACATCAAG 3899
 Qy ---ArgAspValLeuGluProThrGlnGlnGluLeuThrAlaAlaValProValLeu 1226
 Db 3900 ATGGAAGACACTTGAACATACAGATTAAGAGGTGATGAGAAACTGTATCCATTCTG 3959
 Qy AlaLysThrGluValGlyGlnGluGlyGluValAlaAspThrLeuAspGlyLysValLys 1246
 Db 3960 TCAAGAAGTGAAGGGGACCTCA-----GAGGCTGACCAAGTATGCGATGAGAAACCAA 4013
 Qy LysGluSerThrGluValGlnSer-----LeuSerLeuGluGluGlyLys 1294
 Db 4134 GATGATGCTCTTGAACCTGACAGCTCAAGCTCAAGTCTCCATCCCGGTGAGAGAG 4193
 Qy MetGluThrAspValGluLysGluLysArgGluThrLysProGluGlnValSerGluGlu 1314
 Db 4194 ATGTAGTTCAGGTCAAGGAGGAGAAACAGAAAGCAGGCCAACCATGTGATTAAGAG 4253
 Qy Gln---GlnGlnGluThrAlaAlaProGluHisGluGluGluThrTyrGlyLysProValLeu 1333
 Db 4254 AAGCTTGAGACGAAGACAGCTGTACCTGATCTGAAGAGTCAATGACAGCTCTCCAG 4313
 Qy ThrLeuAspMetProSerSerGluArgGlyLysAlaLeuGlySerLeuGlyGlySer--- 1352
 Db 4314 ACAGTGAATGTGCCATCATATGATGGGCAAGAAAGTCAAGCTTTGGAGAAAGCCCT 4373
 Qy ProSerLeuProAspGlnAspLysAlaGlyCysIleGluValGlnValGlnSerLeuAsp 1372
 Db 4374 CCTCCCTGCTAGTCAAGAGGAGGACATAGCAACAAATTTCAAGTCAAGGCTCTGAG 4433
 Qy ThrThrValThrGlnThrAlaGluAlaValAlaGluLysValIle-----GluThrValVal 1390
 Db 4434 GCATATTCACTTAAACAGCGGCTGCAGAGAGAGAAAGTCTTAGAGAAACCTCCAC 4493
 Qy IleSerGluThrGluGluSerProGluCysValGlyAlaHisLeuLeuProAlaGluLys 1410
 Db 4494 ATTTTGAAGACGTTGAACGTTGAGAGCTGCAAGTGCACATTTAGTTCTGGAAGAGAA 4553


```
QY 1411 SerSerAlaThrGlyGlyHisTrpThrLeuGlnHisAlaGluAspThrValProLeuGly 1430
Db 4554 TCCTCTGAAAAAATGAAGACTTTGCGCTCATCCAGGGAAGATGCTGCGCCACAGGG 4613
QY 1431 ProGluSerGlnAlaGluSerIleProIleValThrProAlaProGluSerThrLeu 1450
Db 4614 CCGGACTGTGAGCAAAATCGACACCACTGATAGTATCTCTACTACCAAGAAAGGCTTA 4673
QY 1451 HisProAspLeuGlnGlyGluIleSerAlaSerGlnArgGluArgSerGluGluAsp 1470
Db 4674 AGTTCCGACCTGGAAGAGAGAGAAAACCATCACTGAAGTGAAGTGCAGATGAAGTCAT 4733
QY 1471 LysProAspAlaGlyProAspAlaAspGlyLysGluSerThrAlaIleGluLysValLeu 1490
Db 4734 GAGCAGTTCCTGCCAGGAGGTC-----AAAGTGAGTGTAGCAATTTGAGGATTTA--- 4784
QY 1491 LysAlaGluProGlu-----IleLeuGluLeuGluSerLysSerAsnLysIleValLeu 1508
Db 4785 -----GAGCCTGAAATGGGATTTTGGAACTTGAGACCAAAAGCAGTAACCTTGTCCTCA 4838
QY 1509 AsnValIleGlnThrAlaValAspGlnPheAlaArgThr---GluThrAlaProGluThr 1527
Db 4839 AACATCATCCAGACAGCGTGTACCACTTTGTACGTACAGAGAAACACAGCCACCGAATG 4898
QY 1528 HisAlaTyraSpSerGlnThrGlnValProAlaCysArgLeuAspSerArgGluProAsn 1547
Db 4899 TTGACGTCTGAGTTACAGACACAACTGACGTGATAAAGCTGACAGCCAGCGCTGGA 4958
QY 1548 ArgCysTrpThrLys-----MetLysAspAla-LysMetLysHisProVa 1562
Db 4959 CAGGAACCGAGAAAGAGAGAGAGGAGGAACTCAGCGCTCTGACAGGATGAACACCAATT 5018
QY 1562 lProGlnProArgGlu-----AspLeuGlnValLeuThrValLe 1575
Db 5019 ACTTCAGCCAAAGAGGAGTCAAGTCAACCGCAGTGGGACAGCACACTCTGATATTTC 5078
QY 1575 uGluAlaTrpAlaGlnProArgLysCysLeuProArgLeuGlnLeuLysAlaProVal 1594
Db 5079 AAAGACATGATGTAAGCCTCAGAAAAGACCATGACTGTTGAGGTAGAAGGTTCCACTG 5136

RESULT 4
US-08-973-462-2
; Sequence 2, Application US/08973462B
; Patent No. 6191270
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; APPLICANT: DAUBERSIES, PIERRE
; TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
; FILE REFERENCE: 0660-0125-0 PCT
; CURRENT APPLICATION NUMBER: US/08/973,462B
; CURRENT FILING DATE: 1998-02-06
; EARLIER APPLICATION NUMBER: PCT/FR96/00894
; EARLIER FILING DATE: 1996-06-12
; EARLIER APPLICATION NUMBER: FR 95/07007
; EARLIER FILING DATE: 1995-06-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 5361
; TYPE: DNA
; ORGANISM: P. falciparum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(5361)
US-08-973-462-2

Alignment Scores:
Pred. No.: 8,51e-23 Length: 5361
Score: 462.00 Matches: 323
Percent Similarity: 36.36% Conservative: 293
Best Local Similarity: 19.07% Mismatches: 646
Query Match: 5.72% Indels: 432
DB: 4 Gaps: 68
```

```
US-09-902-432-4 (1-1596) x US-08-973-462-2 (1-5361)
QY 53 GlnLysAsnGlyGlnLeuSerValAsnGlyValAlaGluGlnGlyAspValHisVal 72
Db 283 CAAGTAATGGTGAATAGTAGT-----GAAGAAGTAAGGAAAAAATT 327
QY 73 GlnGluGluAsnGlnGluGlnGlu---GluGluValValAspGluAspValGlyGln 91
Db 328 CTTGACTTATTAGAAAGAAATAATCACTAACTGAAAGTGTAGATGATAATAAAATTTA 387
QY 92 ArgGluSerGluAspValArgGluLys-----AspArgValGluGlu----- 105
Db 388 GAAGAAGCCGAGATATAAGGAAAAATATCTTTAATGAATAATATAGAGAACCAAGAA 447
QY 106 -----MetAlaAlaAsnSerThrAlaValGluAspIle 116
Db 448 AATATTATTGACAAATTTAATAATATATTGGACAAATTCAGAAAAACAAGAAAGTGA 507
QY 117 ThrLysAspGlyGln-----GluGluThrSerGluIleIleGluGlnIleProAla 133
Db 508 TCAGAAATATGTACACAGTCAGTGATGAACCTTTTAAATGAATTTAATAATAGTGTAGATGTT 567
QY 134 SerGluAsnAsnValGluGluMetValGlnProAlaGluSerGlnAlaAsnAspValGly 153
Db 568 AATGGAGAAAGTAAAGAAAAATATTTTGGAG-----GAAAGTCAAGTTAATACGATATT 621
QY 154 PheLysLysValPheLysPheValGlyPheLysPheThrValLysLysAspLysAsnGlu 173
Db 622 TTTAATAGTTTAGTAAAA-----AGTTTCAACAGAACAAACACAC 663
QY 174 LysSerAspThrValGlnLeuLeuThrValLysLysAspGluGlyGluAlaGluAla 193
Db 664 AATGTTGAAGAAAAAGTTGAAGAAAGTGTAGAAGAAATGACCAAGAAAGTGTAGAAGAA 723
QY 194 SerVal-----GlyAlaGlyAspHisGlnGluProSerValGluThrAlaValGlyGlu 211
Db 724 AATGTAGAAGAAATGTAGAAGAAATGACGCGGAAGTGTAGCCTCAAGTGTGTGAAGAA 783
QY 212 Ser-----AlaSerLysGluSerGluLeuLysGlnSerThrGluLysGlnGluGlyThr 229
Db 784 AGTATAGCTTCAAGTGTGATGAAAGTATAGATTCAAGTATTGAAGAAATGTAGTCTCA 843
QY 230 LeuLysGlnGlu----- 233
Db 844 ACTGTTGAAGAAATCGTAGCTCCAAGTGTGTAGAAAGTGTGCGCTCCAAGTGTGAAGAA 903
QY 234 -----GlnSerSerThrGluIleProLeuGlnAlaGluSerAspGlnAlaAlaGluGlu 251
Db 904 AGTGTAGAAGAAAAATGTTGAAGAAAGTGTAGCTGAAAAATGTTGAAGAAAGTGTAGCTGA 963
QY 252 GluAlaLysAspGluGlyGluGluLysGlnGluLysGluProThrLysSerProGluSer 271
Db 964 AATGTTGAAGAAAGTGTAGCTGAAAAATGTTGAAGAAAGTGTAGCTGAAAAATGTTGAAGAA 1023
QY 272 ProSerSerProValAsnSerGluThrThrSerSerPheLysLysPhePheThrHisGly 291
Db 1024 ATCGTAGCTCCAAGTGTGAAGAAATCGTAGCTCCAAGTGTGAAGAAATGTAGTCTCA 1083
QY 292 TrpAlaGlyTrpArgLysLysThrSerPheLysLysSerLysGluAspLeuGlu--- 310
Db 1084 AGTGTGTAGAAGTGTGCGCTCCAAGTGTGAAGAAAGTGTAGAAGAAATGTTGAAGAA 1143
QY 311 ---ThrAlaGluLysArgLysGlnGlnAlaGluLysValAspGluGluGluLysGlu 329
Db 1144 AGTGTAGCTGAAAAATGTTGAAGAAAGTGTAGCTGAAAAATGTTGAAGAAAGTGTAGCTGAA 1203
QY 330 LysThrGluProAla-----SerGluGlnGlnGluProAlaGluAspThrAspGln 346
Db 1204 AATGTTGAAGAAAGTGTAGCTGAAAAATGTTGAAGAAAGTGTAGCTGAAAAATGTTGAAGAA 1263
QY 347 AlaArgLeuSerAlaAspTyrGluLysValGluLeuProLysGlnValGlyAsp 366
Db 347 AlaArgLeuSerAlaAspTyrGluLysValGluLeuProLysGlnValGlyAsp 366
```


[illegible]

Qy	1057	GlusAp	-----SerGluValLeuLeuAspGlu	1061
Db	3055	GAATAATTTATAACAGGTATGTTTCGAAGTATAGAAACCAAGTATAGTAATCCAAATCAGAA	3114	
Qy	1066	LysGluLysAspVal	-----Met	1071
Db	3115	GAAGAAGTGAATTTGAATGAAATGCGGTAGTTCGATTTAGATAATATAGAAATATG	3174	
Qy	1072	ProLysGly	-----ProValGlnGluLeu	1079
Db	3175	AAAGAAGGTTTATTAATAATTTAGAAAAATATTTCAAGTACTGAAGGTGTTCAAGAACT	3234	
Qy	1080	GlyAlaGluHisLeuAlaGlnGly	-----	1087
Db	3235	GTAACCTGAACATGTTAGAACAAATGTATATGCGATGTTGATGTTCTGCTATCAAAAGAT	3294	
Qy	1088	-----SerGluThrGlyGlnAlaThrProGluSerLeuGluValPro	1101	
Db	3295	CAATTTTACGAATATTAATGAGGCAGAGGGTTGAAGAAATATGTTTTTAATTTGGAA	3354	
Qy	1102	GluValThrAlaAspValAspHisValAlaThrCysGlnValIleLys	-----Leu	1118
Db	3355	GATGTTATTAAGAGTGAAGTGAATTAATCTGTGAAGAAATTAAGGATGAACCGGTT	3414	
Qy	1119	GlnGlnLeuMetGluGln	-----AlaValAlaProGluSerSerGluThrLeuThr	1135
Db	3415	CAAAAGAGGTAGAAAAAGAAACCTGTTGATTTATTAAGAAATGGAAGAAATATTTGTA	3474	
Qy	1136	AspSerGluThrAsnGlySerThrProLeuAlaAspSerAspThrAlaAspGlyThrGln	1155	
Db	3475	GATGTTATAGGAAGAAAAAGAAATTTAAACAGACAAG	3531	
Qy	1156	GlnAspGluThrIleAspSerGlnAspSerLysAlaThrAlaValAlaGlnSerGln	1175	
Db	3532	GAATCCATAGAAATATCTTCAGAGTCTTAAGAAGAAACTGAATCTATTAAAGATAAA	3588	
Qy	1176	ValThrGluGluGluAlaThrAlaGlnLysGluGluProSerThrLeuProAsnAsn	1195	
Db	3589	-----CAAAAGAGTGTTCCTAGTTGTTGAAGAA	3618	
Qy	1196	ValProAlaGlnGluGluHisGlyGluGluProGlyValArgAspValLeuGlu	1213	
Db	3619	-----GTTCAAGACAATGATATGGAATGAAGTGTGAGAAGTTTGTAGAAATGGA	3672	
Qy	1214	ThrGlnGlnGluLeuThrAlaAlaValProVal	-----Leu	1226
Db	3673	ATGGAAGAGAGTAAATGAAGTCTGTGTAATAAATGACATTAAGTAACTTATT	3732	
Qy	1227	AlaLysThrGluValGlyGlnGluGlyGluValAlaAspTrpLeu	-----AspGlyGluLysVal	1245
Db	3733	GAAGAAACTCAAGAGTTAAATGAAGTAGAAGCAGATTTAATAAAGATATGGA	3792	
Qy	1246	LysGlu	-----GluGlnGluValPheValHisSerGlyPro	1257
Db	3793	AAAGAATTTAGAAAAAGCATTTATCAGAAGATTTCTAAAGAAATTAATAGATGC	3852	
Qy	1258	AsnSerGlnLysAlaAlaAspValThrTrpAsp	-----SerGluValMet	1272
Db	3853	ACATTAGAAAAAGTTATTGAGAGGNACATGATATAACGACGAGCTTGGATGA	3912	
Qy	1273	GlyValAlaGlyCysGlnGlu	-----	1279
Db	3913	GAATTTAAAGATGTCGAAGAAGACAAGATCGAAAAAGTATCTGATTTAAAGATCTTGA	3972	
Qy	1280	-----LysGluSerThrGluValGlnSerLeuSerLeuGluGlu	1292	
Db	3973	GAAGATATATTAAGAAGAGTAAAGAAATCAAGAACTTGAAGTGAATTTTAGAAGAT	4032	
Qy	1293	GlyGlu	-----MetGluThrAspValGluLysGluLysArgGluThrLysPro	1308
Db	4033	TATAAGAATTAATAAACTATTGAACAGATATTTTAGAAGAGAAAAAGAAATAGAAAA	4092	
Qy	1309	GluGlnValSerGluGlyGluGlnGluThrAlaAlaProGluHisGluGlyThrTyr	1328	

DECLASS

```

US-08-973-462-1
; Sequence 1, Application US/08973462B
; Patent No. 6191270
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; APPLICANT: DAUBESIES, PIERRE
; TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
; FILE REFERENCE: 0660-0123-0 PCT
; CURRENT APPLICATION NUMBER: US/08/973,462B
; CURRENT FILING DATE: 1998-02-06
; EARLIER APPLICATION NUMBER: PCT/FR96/00894
; EARLIER FILING DATE: 1996-06-12
; EARLIER APPLICATION NUMBER: FR 95/07007
; EARLIER FILING DATE: 1995-06-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Patentin ver. 2.0
; SEQ ID NO 1
; LENGTH: 6152
; TYPE: DNA
; ORGANISM: P. falciparum
US-08-973-462-1

```

Alignment Scores:

Pred. No.: 1.02e-22 Length: 6152
 Score: 462.00 Matches: 323
 Percent Similarity: 36.36% Conservative: 293
 Best Local Similarity: 19.07% Mismatches: 646
 Query Match: 5.72% Indels: 432
 DB: 4 Gaps: 68

US-09-902-432-4 (1-1596) x US-08-973-462-1 (1-6152)

Oy 53 GlnlysaenglvglnleuserValaenglyValaIglungllyaspValHisVal 72
 Db 527 CAAGTAATGCGTAATAGCTAGT-----GAAAGAGTAAAGCAAAAAATT 571
 Oy 73 GlnGlnGlnaenglvglnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 91
 Db 572 CTGGACTTATTAAGAAAGAAATACATTACTGAAAGTGAATGATATAATAATTAA 631
 Oy 92 ArgGlnSerGlnaspValargGlnlys-----AspArgValGlnGln----- 105
 Db 632 GAAGAAGCCGAGATATTAAGCAAAATATCTTATTAAGTAATAGAGAACCAAAAGAA 691
 Oy 106 -----MetAlaIaAspSerThrAlaValGlnaspIle 116
 Db 692 AATATTATTGACAAATTATTAAATAATATTGACAAAAATTCAGAAAAACAGAAAGTGA 751
 Oy 117 ThrlysaaspGlyGln-----GlnGlnThrSerGlnIleIleGlnGlnIleProAla 133
 Db 752 TCAGAAAAATCTACAGTCACTGATGAACCTTTTAAGCAATTATTAAATAGTGTAGATT 811
 Oy 134 SerGlnAsnAsnValGlnGlnMetValGlnProAlaGlnSerGlnAlaAsnaspValGly 153
 Db 812 AATGCGAAGTAAAGAAATATTTTGGAG-----GAAAGTCACAGTTATGCGATTATT 865
 Oy 154 PheIysIysValPheIysPheValGlyPheIysPheThrValIlybysaspIysaenglu 173
 Db 866 TTTTAATAGTTTGTATMAA-----AGGTTCAACAGAAACAACAACAC 907
 Oy 174 LysSerAspThrValGlnleuLeuThrValIlybysaspGlnGlyGlnGlnGlnGlnGln 193
 Db 908 AATGTTGAAAGAAAAAGTTGAAGAAAGTGTAGAAAGAAATGACACAAAGTGTAGAGAA 967
 Oy 194 SerVal-----GlyAlaGlyAspHisGlnGlnProSerValGlnThrAlaValGlyGln 211
 Db 968 AATGTGAGAAAGAAATGTAGAAAGAAATGACACGAGAAAGTGTAGCTCAAGTGTGAAGAA 1027
 Oy 212 Ser-----AlaSerIysGlnSerGlnLeuIysGlnSerThrGlnIysGlnGlnGlnGln 229
 Db 1028 AGTATAGCTTCAAGTGTGTATGAAAGTATGATTCAAGTATTGAAAGAAATGTAGCTCCA 1087
 Oy 230 LeuIysGlnGln----- 233
 Db 1088 ACTGTTGAAAGAAATCTAGCTCCAGTGTGTAGAAAGTGTGCTCCAGTGTGAAGAA 1147
 Oy 234 -----GlnSerSerThrGlnIleProLeuGlnAlaGlnSerAspGlnAlaIaGlnGln 251
 Db 1148 AGTGTGAGAAAGAAATGTGAGAAAGTGTAGCTGAAAGAAATGTTGAAAGAAAGTGTAGCTGAA 1207
 Oy 252 GlnAlaIlybysaspGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 271
 Db 1208 AATGTTGAAAGAAAGTGTAGCTGAAAGTGTGAGAAAGTGTAGCTGAAAGTGTGAGAA 1267
 Oy 272 ProSerSerProValAsnSerGlnThrThrSerSerPheIysIysPhePheThrHisGly 291
 Db 1268 ATCGTAGCTCCAGCTGTGAGAAATCGTAGCTCCAGCTGTGAGAAATTTGAGCTCCA 1327
 Oy 292 TrpAlaGlyTrpArgIysIysThrSerPheIysSerIysSerIysaspIysaenglu--- 310
 Db 1328 AGCTGTTGAGAAAGTGTGCTCCAGTGTGAGAAAGTGTAGAGAAATGTTGAGAA 1387
 Oy 311 ---ThrAlaGlnIlybysargIysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 329

Db 1388 AGTGTAGCTGAAAGTGTGAGAAAGTGTAGCTGAAAGTGTGAGAAAGTGTAGCTGAA 1447
 Oy 330 LysThrGlnProAla-----SerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 346
 Db 1448 AATGTTGAGAAAGTGTAGCTGAAAGTGTGAGAAAGTGTAGCTGAAAGTGTGAGAA 1507
 Oy 347 AlaArgLeuSerAlaAspTrpGlnIysValGlnLeuProLeuGlnAspGlnValGlyAsp 366
 Db 1508 ATC---GTAGCTCCAGCTGTTGAAAGAAATCGTAGCTCCAGCTGTTGAAAGAAATT- 1558
 Oy 367 LeuGlnAlaSerSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 386
 Db 1559 GTAGCTCCAGCTGTTGAGAAAGTGTGCTCCAGCTGTGAGAAAGTGTGAGAAAGTGTGAGAA 1618
 Oy 387 MetGlnAlaHisGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 406
 Db 1619 GTTGA-----GAAAGTGTAGCTGAAAGTGTGAGAAAGTGTGAGCTGAAAGTGTGAA 1672
 Oy 407 Gln 426
 Db 1673 GAAAGT-----GTAGCTGAAAGTGTGAGAA 1699
 Oy 427 SerLeuProProGlnIlybysValGlnProGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 446
 Db 1700 ACTGTA-----GTGAAAGTGTGAGAA 1717
 Oy 447 GlnGlnLeuMetLysSerArgGlnMetCysValSerGlyGlyAspHisThrGlnLeuThr 466
 Db 1718 GAAGAAAGTGTAGCTGAA----- 1735
 Oy 467 AspLeuSerProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 486
 Db 1736 -----AATGTTGAGAAAGTGTAGCTGAAAGTGTGAGAAAGTGTGAGAAAGTGT 1789
 Oy 487 GlnMetLeuSerSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 506
 Db 1790 GAAGAAATCGTAGCTCCAGCTGTGAGAAAGTGTGAGAAAGTGTGAGAAAGTGTGAGAA 1849
 Oy 507 SerSerSerGlyLeuLysLysLeuSerGlyLysLysGlnIysGlyLysArgGlyGlyGly 526
 Db 1850 GCTCCAGTGTGTGAGAAAGTGTGCT----- 1876
 Oy 527 GlyAspGlnGlnProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 546
 Db 1877 -----CCAGTGTGAGAAAGTGTGAGAAAGTGTGAGAAAGTGTGAGAAAGTGTGAGCT 1924
 Oy 547 GlnGlnIysGlyGlnSerSerAlaSerSerProGlnGlnPro-----Gln 561
 Db 1925 GAAATGTTGAGAAAGTGTAGCTGAAAGTGTGAGAAAGTGTAGCTGAAAGTGTGAA 1984
 Oy 562 GlnThrThrCysLeuGlnIlybysGlyProLeuGlnAlaProGlnAspGlyGlnAlaGlnGln 581
 Db 1985 GAAAGTGTAGCTGAAAGTGTGAGAAATCGTAGCTCCAGCTGTGAGAAATCGTAGCT 2044
 Oy 582 GlyThrThrSerAspGlyGlnIlybysArgGlnGlyIleThrProTrpAlaSerPheIys 601
 Db 2045 CCAGCTGTT-----GAGAAATGTAGCTCCA-----AGTGTGTA 2080
 Oy 602 LysMetValThrProLysLysArgValArgArgProSerGlnSerAspLysGlnGlnGln 621
 Db 2081 GAAAGTGTGCTCCAGT-----GTTGAGAAAGTGTGAGAAAGTGTGAGAAAGTGTGAGAAAGT 2134
 Oy 622 Leu---GlnLysValIlybysSerAlaThrLeuSerThrAspSerThrValSerGlnMet 640
 Db 2135 GTAGCTGAAAGTGTGAGAAAGTGTAGCTGAAAGTGTGAGAAAGTGTAGCTGAAAGT 2194
 Oy 641 GlnAspGlnVal-----LysThrValGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 658
 Db 2195 GTTGAAGAAATGTAGCTCCAGCTGTGAGAAAGTGTAGCTCCA----- 2239
 Oy 659 ArgValAspThrSerValSerTrpGlnAlaLeuIleCysValGlySerSerIlybysArg 678
 Db 2240 -----ACTGTTGAGAAATTTGAGCTCCAGCTGTGAGAAAGTGTG 2281

Qy	679	Ala	Arg	Lys	Ala	Ser	Ser	Asp	Asp	Glu	Gly	Gly	Pro	Arg	Thr	Leu	Gly	Cys	Asp	Ser	698
Db	2882	GCT	CCA	AGT	GTT	GAA	GAAG	AGT	GTA	GAAG	AAAT	GTT	GTA	GAAG	AAAGT	GTGA					2329
Qy	699	His	Arg	Ala	Glu	Glu	Ala	Ser	Lys	Asp	Lys	Glu	Ala	Gly	Thr	Asp	Ala	Val	Pro	Ala	718
Db	2330	-----	GCT	GAA	AAT	GTT	GAA	GAAG	AGT	GTA	GCT	GAA	AAT	GTT	GTA	GAAG	AAAGT	GTGA			2374
Qy	719	Thr	Gln	Glu	Gln	Asp	Gln	Ala	Gln	Gly	Ser	Ser	Pro	Glu	Pro	Ala	Gly	Ser	Pro	Ser	738
Db	2375	GTAG	CTG	AAA	AAT	GTT	GAA	GAAG	AGT	GTA	GCT	GAAA	ATG	TGT	GAA	GAAT	CGT	AGT	CCA	ACT	2434
Qy	739	Glu	Gly	Glu	Gly	Val	Ser	Thr	Trp	Glu	Ser	Phe	Lys	Arg	Leu	Val	Thr	Pro			755
Db	2435	GTT	GAA	GAAT	TCG	TAG	CT	-----	CCA	ACT	GTT	GTT	GAA	GAAT	TGT	GAC	TCC	CAAG	TGT	TGA	2488
Qy	756	-----	Arg	Lys	Lys	Ser	Lys	Leu	Glu	Glu	Lys	Ala	Glu	Asp	Ser	Ser	Val				772
Db	2489	GAA	AGT	TGC	GCT	CCA	AGT	GTT	GAA	GAAG	AGT	GTA	GAAG	AAAT	GTT	GAA	GAAG	TGT	TAG	CT	2548
Qy	773	Glu	Gln	Leu	Ser	Thr	Glu	Leu	Glu	Pro	Ser	Arg	Glu	Ser	Trp	Val	Ser	Leu	Lys		792
Db	2549	GAAT	CTG	TCA	GAA	AGT	GTA	GCT	GAAA	ATG	TGT	GAA	GAAGT	-----							2590
Qy	793	Phe	Leu	Pro	Gly	Arg	Arg	Lys	Lys	Arg	Ala	Asp	Gly	Lys	Gln	Glu	Gln	Ala			810
Db	2591	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	
Qy	811	Val	Glu	Asp	Ser	Gly	Pro	Val	Glu	Leu	Asn	Glu	Asp	Asp	Pro	Asn	Val	Pro	Ala	Val	830
Db	2627	GTT	GAA	GAAT	TGT	TAG	CTC	CAAG	TGT	GAA	GAA	-----	-----	-----	-----	-----	-----	-----	-----	-----	2668
														</							

Qy	1022	LeuGlnAlaValAlaAspLysValLysGluGluSerGlnValProAlaThrGlnThrVal	104
Db	3179	TTTAATACTGTATAGATAAAGTAGAGAAACAGTAGAAATTTACGGGAAAGTTTGA	3238
Qy	1042	GlnArgThrGlySerLysAla-----LeuGluLysValGluGluValGlu	1056
Db	3239	AACAATGAAATGATAAAGCATTTTTTTAGTGAAATATTTTGATATGTAAAGGAATACAA	3298
Qy	1057	GluAsp-----SerGluValLeuAlaSerGlu	1065
Db	3299	GAAATTTTAAACAGGTATGTTTCGAAGTATAGAAACCAAGTATAGTAAATCCATCAGAA	3358
Qy	1066	LysGluLysAspVal-----Met	1071
Db	3359	GAAAGAGTTGATTTGGAATGAAATGCGTTAGTTCGATTTAGTATAATATAGAAATATG	3418
Qy	1072	ProLysGly-----ProValGlnGluAla	1079
Db	3419	AAAGAAGTTTATTAATAATAATAGAAATATTTTCAAGTACTGAAGGTGTTCAAGAACT	3478
Qy	1080	GlyAlaGluHisLeuAlaGlnGly-----	1087
Db	3479	GTAACACATAGTAAACAAATGTATATGCGATGTGATGTTGCTCTGCTATGAAAGAT	3538
Qy	1088	-----SerGluThrGlyGlnAlaThrProGluSerLeuGluValPro	1101
Db	3539	CAATTTTATAGGAATATTAATAGGCGAGGAGGTTGAAAGAAATGTTTTTAAATTTTCGAA	3598
Qy	1102	GluValThrAlaAspValAspHisValAlaThrCysGlnValIleLys-----Leu	1118
Db	3599	GATGTATTTAAAGTGAAGTGAATGTAATCTGTGAAGAAATTAAGGATGAACCGGTT	3658
Qy	1119	GlnGlnLeuMetGluGln-----AlaValAlaProGluSerSerGluThrLeuThr	1135
Db	3659	CAAAAGAGCTAGAAAAGAAACCTGTACTATTATTGAAGAAATCGAAGAAATATTGTA	3718
Qy	1136	AspSerGluThrAsnGlySerThrProLeuAlaAspSerAspThrAlaAspGlyThrGln	1155
Db	3719	GATGTATTAGAGGAAGAAAAGAAAGATTAAACAGACAAAG---ATCATAGATGCGATAGAA	3775
Qy	1156	GlnAspGluThrIleAspSerGlnAspSerLysAlaThrAlaAlaValAlaGlnSerGln	1175
Db	3776	GAATCCATAGAAATATCTTCAGATTTAAAGAAAGAACTGAATCTATTAAAGATAGAAA-	3832
Qy	1176	ValThrGluGluGluAlaAlaThrAlaGlnLysGluGluProSerThrLeuProAsnAsn	1195
Db	3833	-----GAAAAGATGTTTCACCTAGTTGTTGAAGAA-----	3862
Qy	1196	ValProAlaGlnGluGluHisGlyGluGluProGlyAspValLeuGlu-----Pro	1213
Db	3863	-----GTTCAAGACAATGATATGGAAGTGTGTAGAAAGTTTTAGAAGTTTGAATAAT	3916
Qy	1214	ThrGlnGlnGluLeuThrAlaAlaValProVal-----Leu	1226
Db	3917	ATGAAGAGAGAGTAAATGAAGATGCTGTGTGAATTAATATGACATTACTAGCAAACTTATT	3976
Qy	1227	AlaLysThrGluValGlyGlnGluGlyGluValAspTrpLeu---AspGlyGluLysVal	1245
Db	3977	GAAGAACTCAAGGTTAAATGAGTAGAAGCAGAGATTAAATAAAGATATGCAAAAATTA	4036
Qy	1246	LysGlu-----GluGlnGluValPheValHisSerGlyPro	1257
Db	4037	AAAGAAATAGAAAAAGCATTATCAGAAGATTTCTAAAGAAATAATAGATGCAAAAAGATGAT	4096
Qy	1258	AsnSerGlnLysAlaAlaAspValThrTyrAsp-----SerGluValMet	1272
Db	4097	ACATTAGAAAAGATTATTGAAGAGCAACATGATATATACGACGACGTTGGATGAAGTTGTA	4156
Qy	1273	GlyValAlaGlyCysGlnGlu-----	1279
Db	4157	GAATTTAAAGATGTCGAAGAGCAAGATCGAAAAGATATCTGATTAAAGATCTTGAA	4216
Qy	1280	-----LysGluSerThrGluValGlnSerLeuSerLeuGluGlu	1292

```

Db 4217 GAAGTATATTAAAGAACTAAAGAAATCAAGAACTTGAAGTGAATTTAGAAAGT 4276
Oy 1293 Glylu-----MetgluthraspValGluYsGluYsArgGluThryPro 1308
Db 4277 TTTAAAGAAATTAAGAACTTTGAAACAGATATTTTGAAGAGAAAAAGAAATAGAAAA 4336
Oy 1309 GluGlnValSerGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 1328
Db 4337 GATCATTTTGAAGAAATTCGAAGAGAGCTGAAGAAATTAAGATCTTGAACAGATATA 4396
Oy 1329 GlyYsProValLeuThryLeuaspMetProSerSerGlu----- 1341
Db 4397 TTTAAAGAAATCTCTCATTTAGAGTTGAAGAAAAAAATTTAGAAAGATACGAA 4456
Oy 1341 ----- 1341
Db 4457 TTTAAAGAAAGAGTAGAACATTAATAGTGGTATCGCATATTAAGGTTTGAAGAA 4516
Oy 1342 -----ArgGlyYsAlaLeuGlySerLeuGlyGlySer 1352
Db 4517 GATGATTTAGAAAGAGTAGATGATTTTAAAGAGAGTATTTAGACATGTTAAAGGAGAT 4576
Oy 1353 ProSerLeuProaspGlnaspGlnaspGlnaspGlnaspGlnaspGlnaspGlnasp 1372
Db 4577 ATGGAATTTAGGGGATATGATTAAG-----GAAAGTTTGA 4612
Oy 1373 ThrThrValThrGlnThrAlaGluAlaValGluYsValIleGluThryValIleSer 1392
Db 4613 GATGTAACAACAACAACTTGAGAGAAAGATTGAATCC---TTAAAGATGTTTATCTACT 4669
Oy 1393 GluThryGlySerProGluYsValGlyAlaHis----- 1404
Db 4670 GCATTGCGCATGATGAGAAACAATGAAACAAAGAAAAAAAGCTCAAGACCTTAAGTTG 4729
Oy 1405 -----LeuLeu-----ProAlaGluYsSerSerAlaThr 1414
Db 4730 GAAGAAGATTTATTAAAGAGAGGTTAAAGAAAGAAACAAGAAAAAATTAACAAAAAAG 4789
Oy 1415 GlyGlyHisThrThrLeuGlnHisAlaGlu-----AspThrValProLeuGlyProGlu 1432
Db 4790 AAGTAAAGGTTTATTTAAGGATAGGAAACCAAAAGATGAATA---GTAGAAAGTTGA 4846
Oy 1433 SerGlnAlaGluSerIleProIleIleValThrProAlaProGluSerThrLeuHisPro 1452
Db 4847 ATGAAGATGAGAGATATA-----GAAAGAGATGAGAGAA 4882
Oy 1453 AspleuGlnGlyGluIleSerAlaSerGlnArgGluYsSerGluGluGluGluGluGlu 1472
Db 4883 GATATGAAAGAGATATTAAGAAAGATTAAGTTGAAGATATAGATGATAGATGA 4942
Oy 1473 AspAlaGlyProaspAlaaspGlyYsGluSerThrAlaIleGluYsValLeuYsAla 1492
Db 4943 GATATAGGTGAACAAAGAT-----GAAGTATAGATTTAATAGTCCAAAA 4990
Oy 1493 GluProGluIleLeuGluLeuGluSerYsSerAsnYsIle 1506
Db 4991 GAGAAACGCATTGAAAGGTTAAAGCAAAAAAGAAAAAATTA 5032

RESULT 6
US-08-119-125A-2
; Sequence 2, Application US/08119125A
; Patent No. 5610011
; GENERAL INFORMATION:
; APPLICANT: SMITH, Hilda Elizabeth
; APPLICANT: VECITT, Uri
; TITLE OF INVENTION: DNA Sequences which code for Virulence
; TITLE OF INVENTION: Characteristics of Streptococcus suis and parts thereof, poly
; TITLE OF INVENTION: Antibodies derived therefrom and the use thereof for the diagn
; TITLE OF INVENTION: protection against infection by S. suis in mammals, including
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESS: Central Diegeneskundig Instituut

```

```

STREET: Edelhertweg 15
CITY: PH Leystad
STATE:
COUNTRY: The Netherlands
ZIP: NL-8219
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS v.6.0
SOFTWARE: WordPerfect v. 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/119,125A
FILING DATE: 20-SEP-1993
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/NL92/00054
FILING DATE: 19-MAR-1992
APPLICATION NUMBER: NL 9100510
FILING DATE: 21-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Handal, Anthony H.
REGISTRATION NUMBER: 26275
REFERENCE/DOCKET NUMBER: SMITHHE119125
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203) 838-8589
TELEFAX: (203) 838-8794
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 6744 base pairs
TYPE: Nucleic acid with corresponding amino acids
STRANDEDNESS: single stranded
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA
ORIGINAL SOURCE:
ORGANISM: Streptococcus suis type II (pathogenic)
FEATURE:
OTHER INFORMATION: Extracellular factor related protein (EF*) gene
FEATURE:
NAME/KEY: promoter -35 region
LOCATION: bp 66 to 71
FEATURE:
NAME/KEY: promoter -10 region
LOCATION: bp 89 to 94
FEATURE:
NAME/KEY: promoter -35 region
LOCATION: bp 153 to 158
FEATURE:
NAME/KEY: promoter -10 region
LOCATION: bp 176 to 181
FEATURE:
NAME/KEY: ribosome binding site
LOCATION: bp 350 to 356
FEATURE:
NAME/KEY: signal peptide
LOCATION: bp 361 to 498
FEATURE:
NAME/KEY: start of repetitive units R1-R11
LOCATION: bp 2869, 3097, 3292, 3520, 4087, 4381, 4609, 4837,
5065, 5293, 5521.
FEATURE:
NAME/KEY: start of repetitive Asn-Pro-Asn-Leu sequences
LOCATION: bp 2932, 3160, 3355, 3583, 4150, 4444, 4672, 4900,
5128, 5356, 5584.
FEATURE:
NAME/KEY: dyad symmetry regions
LOCATION: from bp 6554 to 6566 and from bp 6571 to 6583
FEATURE:
NAME/KEY: dyad symmetry regions
LOCATION: from bp 6611 to 6625 and from bp 6631 to 6644
US-08-119-125A-2
Alignment Scores: 7.62e-20 Length: 6744
Pred. No.:

```

Score: 421.50 Matches: 333
 Percent Similarity: 36.25% Conservative: 268
 Best Local Similarity: 20.08% Mismatches: 664
 Query Match: 5.22% Indels: 393
 DB: 1 Caps: 71

US-09-902-432-4 (1-1596) x US-08-119-125A-2 (1-6744)

QY 4 GlySerSerThrGluGlnArgSerProGluGlnProAlaGly----- 17
 Db 2056 GGTACAAACACTAAGATTACTTCAGAACTCGATCTACGGGAGCAACTGCAAAACCAAGGA 2115
 QY 18 -----SerAspThrProSerGluLeuValLeuSerGlyHisGlyProAlaAla 33
 Db 2116 GATGACGGTCAATCTTCAACTAAGTTTAAGATTACAGTACCGGACCTGCTTACA 2175
 QY 34 GluAlaSerGlyAlaAlaGlyAspProAlaAspAlaAspProAlaThrLysLeuProGln 53
 Db 2176 GAAGGTACCGGCACTTAT----- 2193
 QY 54 LysAsnGlyGlnLeuSerSerValAsnGlyValAlaGluGlnGlyAspValHisValGln 73
 Db 2194 -----AAGCTTCGGTGTGA 2208
 QY 74 GluGluAsnGlnGlu-----GlyGlnGluGluGluValValAsp-----GluAsp 88
 Db 2209 GAAGATAAAGTATCTTTGTGTCAGAGGGGAAACTTGTGATGGAATAAACCAGCAAAAT 2268
 QY 89 ValGlyGlnArgGluSerGluAspValArgGluLysAspArgValGluGluMetAlaAla 108
 Db 2269 GTAGTTTGACATCTGTAAAGTTACCTTCGTAAA-----CATGCTACGGTGTCA 2319
 QY 109 AsnSerThrAlaValGluAspIleThrLysAspGlyGlnGluGluThrSerGluIle 128
 Db 2320 ACACCACTTCTGTTGTAATCCAGCTAACTTAACGCCAGAGAAAGCCGCGAGTTATT 2379
 QY 129 GluGlnIleProAlaSerGluAsnAsnValGluGluMetValGlnProAlaGluSerGln 148
 Db 2380 GCTCAAAATCAAGAAAGAACACGACGACCAACAGAGATTGAAGGGCTTGCAGATTTCA--- 2436
 QY 149 AlaAsnAspValGlyPheLysLysValPheLysPheValGlyPheLysPheThrValLys 168
 Db 2437 -----GCATTTCAGTTTAAAC 2451
 QY 169 LysAspLysAsnGluLysSerAspThrValGlnLeuLeuThrValLysLysAspGluGly 188
 Db 2452 TCAGATGGT-----ACTGTGTCAGTTGACTACAGT 2481
 QY 189 GluGlyAlaGluAlaSerValGlyAlaGlyAspHisGlnGluProSerValGluThrAla 208
 Db 2482 GCCGGTGGTGTCAATGTTGATGGTGCACACACATTTAAGAAATGCTACCAAACTTG 2541
 QY 209 ValGly-----GluSerAlaSerLysGluSerGluLeuLysGlnSerThrGluLysGln 226
 Db 2542 GCAGATACAGGAATGAAGCAAGCAAGCAATCGACAAATAGCTGAAACATAAATAA 2601
 QY 227 GluGlyThrLeuLysGlnGluGlnSer-----SerThrGluIleProLeuGln 242
 Db 2602 GCTATCGAAGCAAAACGGGATGAAGCGTTTCTAAATTTGATGATGACATTTCTTTGAGA 2661
 QY 243 AlaGluSerAspGlnAlaAlaGluGluAlaLysAspGluGlyGluGluLysGlnGlu 262
 Db 2662 GCAGAACAGACAGAGCGTGTCAAGGATTAAGGATGCTGCGGGATGCTTTGAAA 2721
 QY 263 LysGluProThrLysSerProGluSerProSerProValAsnSerGluThrThrSer 282
 Db 2722 GAATTAGAACAAAGCGACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 2781
 QY 283 SerPheLysLysPheThrHisGlyTTPAlaGlyTTPArgLys----- 297
 Db 2782 GAAATCAATGATCTAAGACTAATGGTTCAGATTAACTGACAGTGCAGAGCAAGCAAGTACG 2841
 QY 298 LysThrSerPheLysLysSerLysGluAspLeuGluThrAlaGlu-----LysArgLys 316

Db 2842 GAAAAAGCTATTAAACAGCGAGGAAAGAACTGCGCAAAAGCAAGATTGAAAAA 2901
 QY 317 GluGlnGluAla---GluLysValAsp-----GluGluGlu----- 327
 Db 2902 GCATTCGAGGCAATGGAAAAAGTTAACAATAATCAAACTTGTGTAAGAGAGAGAAAAA 2961
 QY 328 -----LysGluLysThrGluProAlaSerGluGluGlnGluProAla 341
 Db 2962 GCATACTTTTGATGATATTAAAGAACTTAAGAAAGTTGCAAGTTGAGAAAAATCAATAATGCT 3021
 QY 342 GluAspThr-----AspGlnAlaArgLeuSerAlaAspTyrGlu 354
 Db 3022 GAAATACTGCTGAAATTCGCGAGCAATTCAGAAAGCGAAATTCATACATGAAGAT 3081
 QY 355 LysValGluLeuProLeuGluAspGlnValGlyAspLeuAlaSerGluGluLys 374
 Db 3082 GTTATTATACGACGCCCACTTGTATGCTTTGCAATAAGCTTGAAAAAGATAGCGAAGAACT 3141
 QY 375 CysAlaProLeuAlaThrGluValPheAspGluLysMetGluAlaHisGlnGluValVal 394
 Db 3142 AAGCCAGCTATT-----GATGCTAATCCAACTTAACT 3174
 QY 395 AlaGluValHisValSerThrValGluLysThrGluGluGlnGlnGlyGlyGlyGlu 414
 Db 3175 CCGGAAGAGAAAGCGAAAGCTATTGCTAAGGTAGAGAGCTGTT-----AATAAT 3225
 QY 415 AlaGluGlyGlyValValGluGlyThrGlyGluSerLeuProGluLysLeuAla 434
 Db 3226 GCTGATCTGACATTTTGTGCAAGCTCCCGAAGAACAGTT----- 3267
 QY 435 GluProGlnGluValProGlnGluAlaGluProAlaGluGluLeuMetLysSerArgGlu 454
 Db 3268 -----CAAGCAGTGGAGGATAAGGCTGAC-----AAAGATCTTGCCAAAGTAGAACTT 3315
 QY 455 MetCysValSerGlyGlyAspHisThrGlnLeu-----ThrAspLeuSerProGlu 471
 Db 3316 CAAGCAGCAGCAGCAGCGTGGAGAAAGCATTTGAAGCAAAATCCGAATTTGATCCAGAA 3375
 QY 472 GluLysThrLeuProLysHisProGluGlyIleValSerGluValGluMetLeuSer 491
 Db 3376 GAGAAAGATGATAGCTAAGAAG-----GCAGTA 3402
 QY 492 GlnGluArgIleLysValGlnGlySerProLeuLysLys----- 504
 Db 3403 GAAGACGCGGTTAAGGTGGCGCAGACGCTATTGTAAGCGCTCAACTCCAACCGAAGTT 3462
 QY 505 ---LeuPheSerSerSerGlyLeuLysLysLeuSerGlyLysLysGlnLysArg 523
 Db 3463 GACACAGCAGCAGCAGCGATGGAGGCTATTGATGCGAAGAGAGTTTAAAGCTACTCAG 3522
 QY 524 GlyGlyGlyGlyAsp-----GluGluProGlyGluTyGlyHisIleHisThrGlu 540
 Db 3523 AAGATGCTAAGAACAAAGATTGCCAAAGAACGAAATCAGCTAAGAAAGGATTGACGAC 3582
 QY 541 SerProGluSerAlaAspGluGlnLysGlyGluSerSerAlaSerSerProGluGluPro 560
 Db 3583 AATCCAAACTTGACTCCAGATGAGAAG---GAATCAGCTAAGAAATCGCAGTGAAGAGCG 3639
 QY 561 GluGlu-----ThrThrCysLeuGluLysGlyProLeuGluAlaProGlnAspGlyGlu 578
 Db 3640 GCTAAGGTAGCAACAGCGCTATTGATAAGCA-----TCAACTCCAGATGCAAGTCA 3693
 QY 579 AlaGluGluGlyThrThrSerAspGlyGluLysLysArgGluGlyIleThrProTTPAla 598
 Db 3694 GTAGAAGAG-----GACAAAGGTGAGCAGCTATCAATTTGATTCTTCCCAAGGCA 3744
 QY 599 SerPheLysLysMetValThrProLysLysArgValArgProSerGluSerAspLys 618
 Db 3745 GATGCTAAAGGTGCTCATTCGCTAAGTTGGCAGATGAAATCAAGAAAGCTCGAAGATAAG 3804
 QY 619 GluGluGluGluLys---ValLysSerAlaThrLeuSerSerThrAspSerThrVal 637
 Db 619 ---ValLysSerAlaThrLeuSerSerThrAspSerThrVal 637

Db 3805 CAAGCAGAAAGCAAAAGCTATCGATGGTCACATGACTAATGAGGAAAGCAATC 3864
 QY 638 SerGlu-----MecGlnAspGluValLysThrValGlyGluGluLysProGlu--- 654
 Db 3865 GCTAAGAAAGGCTCTTCATGATGTTGTA-----GATTAAGCAAAAGCAGAGCTT 3912
 QY 655 GluProLysArgValAspThrSerValSerTrpGluAlaLeuLeuValGlySer 674
 Db 3913 GAAGACGACGCTTAAGGATCAACAATGAGATTCAATGAGCT-----ACT 3957
 QY 675 SerLysArgAlaArgLysAlaSerSerSerAspAspGluGlyGlyProArgThrLeu 694
 Db 3958 ACTACAGAAAGAAAGCGAAGCGGGAACCTTGCTGGGGAAGAAAGCTTGACAGACACA 4014
 QY 695 GlyGlyAspSerHisArgAlaGluGluAlaSerLysAspLysGluAlaGlyThrAspAla 714
 Db 4015 GGTAAAGAGAGCTAGAGATCAGTTGAAATGGCTAAGATTAAGAAATTACCTAAGAGAGA 4074
 QY 715 ValProAlaSerThrGlnGlnGlnAspGlnAlaGlnGlySerSerSerProGluProAla 734
 Db 4075 ATCCGAACGAAAGAAAGAAAGCTACTAAATAGTAGAAGAACTTGACAGAGATACGCC 4134
 QY 735 GlySerProSerGluGlyGluGlyValSerThrTrpGlu-----SerPheLysArg 751
 Db 4135 AAAGCTATGAGAGACATCCAAACTGTCAAGATGAAGATTAAGCAAGCGGAATTAAAG 4194
 QY 752 LeuValThrProArgLysLysSerLysSerLysLeuGluGluLysAlaGluAspSerSer 771
 Db 4195 CTAACTGACGCTGTGGCAAAACTTTCAGACACATTCGTGACATCAATGAGTAAAGCTAGC 4254
 QY 772 ValGlu-----GlnLeuSerThrGlnLeuGluProSerArgGluGluSerTrp 787
 Db 4255 CAAAGAAAGCAAAAGCTCAAGCCCTTACGACAGATCTTGAAAGAAAGCTTAAGAAACAGAA 4314
 QY 788 ValSerLysLysLysPheLeuProGlyArgArgLysLysArgAlaAspGlyLysGlnGlu 807
 Db 4315 ATTGCGATTAAGAGCTCCGATGTAGTGTGACATCTTGTAAGATGTGTAGCTTGA 4374
 QY 808 GlnAlaThrValGluAspSerGlyProValGluLeuAsnGluAspAspProAsnValPro 827
 Db 4375 ---GCTACTAAACAGAT----- 4389
 QY 828 AlaValAlaProLeuSerGluTrpAsnAlaValGluArgGluLysMetGluAlaGlnGly 847
 Db 4390 GCTAAGAACAGATTTGCTAAAGATGCAAGCGCTGCTTAAGAACCAATTGCAAGCAATCCA 4449
 QY 848 AsnThrGluLeuProGlnLeuGluAlaValLysValSerGluGluLeuSerLysThr 867
 Db 4450 AAC-----TTGACAGACCCAGAGAAAGAAC 4476
 QY 868 LeuValHisThrValSerValAlaValIle-----AspGlyThrArgAlaValThr 884
 Db 4477 TTCACCGATGCGGTAGATGACAGAAAGTACCAAGCTAACGACCAATTTCAGCTGCAACC 4536
 QY 885 Ser-----ValGluGluLysSerProSerTrpLysSerAlaSerValThrGluPro 901
 Db 4537 AGCCGACGAGATGTTCAAAAAGAGAGATGCGGTGTCACCAATTCGACAGAAAGTGT 4596
 QY 902 LeuGlnHisThrAlaGlyLysLysMetProProVal-----GluGluValThrGluLys 919
 Db 4597 CTTGACGCGACTTAACAAAGATTAAGATTAAGATTGCTTAAGATGCAAGCGCTGTAA 4656
 QY 920 AspLeuIleAlaGluGluThrProValLeuThrGlnThrLeuProGluGlyLysAspAla 939
 Db 4657 GAAGCAATTGGC---TCCATCCAACTTGACAGACGA-----GAGAAGAAAGAAC 4704
 QY 940 HisAspAspMetValThrSerGluValAspPheThrSerGluAlaValThr---AlaThr 958
 Db 4705 TTCAACGAGTGGGTATGATGACAGAAAGTACCAAGCTAACGCAATTTTCAGCTGCAACC 4764
 QY 959 GluThrSerGluAlaLeuArgThrGlnGluValThrGluAlaSerGlyAlaGluGluThr 978
 Db 4765 AGCCCAAGCATGTTCAAAAAGAGAGATGACGAGTGTTCACCACTTGACAGAAAGATGT 4824

QY 979 ThrAspMetVal-----SerAlaValSerGlnLeuThrAspSerProAsp 993
 Db 4825 CTTGACGAGCTAAACAAAGATGCTAAGATTAAGATGCTTAAGAAATGCGACGCTGAAG 4884
 QY 994 ThrThrGluGluAlaThrProValGlnGluValGluSerGlyValLeuAspThrGluGlu 1013
 Db 4885 TCAGGCATTGACGCAATCCA-----ACTTGACAGATGACAGAGAG 4926
 QY 1014 Glu-----GluArgGlnThrGlnAlaIleLeuGlnAlaValAlaAspLysValLysGlu 1031
 Db 4927 GATACAGCTAAGAAAGCAGTTGATGCTGATGCTTAAGCTCGACAGATGATGATGACT 4986
 QY 1032 GluSerGlnValProAlaThrGlnThrValGlnArgThrGlySerLysAlaLeuGluLys 1051
 Db 4987 TCACACATCCAGTCGAAAGCCCAATCGGCAAGAGACAAAGGC----- 5028
 QY 1052 ValGluGluValGluGluAspSerGluValLeuAlaSerGluLysGluLysAspValMet 1071
 Db 5029 GTAGGTCATTCGCCCAAGATCTTTCAGCCAGCGAAACAGATGCTTAAGACAAAGATT 5088
 QY 1072 ProLysGlyProValGlnGluAlaGlyAlaGluHisLeuAlaGlnGlySerGluThrGly 1091
 Db 5089 -----GCCAAAGAAAGTTGCCGCGACGCTAAAGAACCAATT 5121
 QY 1092 GlnAlaThrProGlu-----SerLeuGluValProGluValThrAlaAspVal 1107
 Db 5122 GATGCCAATCCGAACCTTATCAAGATGACAGAGAAAGAGCTTCAAGAAAGCGTGAATCA 5181
 QY 1108 AspHisValAlaThrCysGlnValIleLysLeuGlnGlnLeuMetGluGlnAlaValAla 1127
 Db 5182 GATGCTTAAGCTACGACAGATGCAATT-----GATGCTTCAACCAAGT 5223
 QY 1128 ProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrProLeuAlaAsp 1147
 Db 5224 CCAGTCGAAGCGCAATCG-----GCAGAGGACAAAGCGGTGATTCATCGCC--- 5271
 QY 1148 SerAspThrAlaAspGlyThrGlnGlnAsp-----GluThrIleAspSerGlnAsp 1164
 Db 5272 CAAGATGTTCTTGAGCGACGCAAGCAAGATGCTTAAGATTAAGATTGCTTAAGAAATCCAG 5331
 QY 1165 SerLysAlaThrAlaAlaValArgGlnSerGlnValThrGlnGluGluAlaAlaThrAla 1184
 Db 5332 GCTGCTAAGTCAAGCATTTGACCCGAATCCAAACTTGACAGATGACAGAAAGAAATCACT 5391
 QY 1185 GlnLysGlu-----GluProSerThrLeu 1192
 Db 5392 AAGAAAGCGGTAGATGACAGATGCTTAAGCTGCGACAGATGCAATGATGATTCACAAAGT 5451
 QY 1193 ProAsnAsnValProAlaGlnGluGluHisGlyGluGluPro---GlyArgAspValLeu 1211
 Db 5452 CCAGTCGAAGCGCAATCGGCAAGAGCAAGAGCGCTAGGCGCCATCGCCAAAGACATTTCT 5511
 QY 1212 GluProThrGlnGlnGlnLeuThrAlaAlaValProValLeuAlaLysThrGluVal 1231
 Db 5512 GATGCCCGGCAACAGAT-----GCTTAAGAACAGATT 5544
 QY 1233 GlyGlnGluGluValAspTrpLeuAspGlyLysValLysGluGlnGluVal 1251
 Db 5545 GCTAAAGAGGCA-----GAATCCGCTAAGTCA 5571
 QY 1252 PheValHisSerGlyProAsnSerGlnLysAlaAlaAspValThrTyraAspSerGlu--- 1270
 Db 5572 GTCAATTGACTCCAAATCCAACTTGACAGATGAGCTTAAGAGAGCGGCTTAATCTAAATT 5631
 QY 1271 -----ValMetGlyValAlaGlyCysGlnGlnGluLys 1280
 Db 5632 CATTAAGCTGTTGAGAGAGCAATTTTAAATCAATGATGTTTGAACCTTATCAAGAGTTG 5691
 QY 1281 GluSerThrGluValGlnSerLeuSerLeu----- 1290
 Db 5692 GAAAAATCAAACTTCCAATGCGACGCTTAATTAAACAGCTGCGAAAGTAAACACACAGTG 5751

Oy	141	ctvalgnProalagluserctlnlaashnspvalGlyPheylsValPheylsPheva	161
Db	1530	CAATGCCACCATCCGACGAAAGGATCAAGAACTG-----GC	1565
Oy	161	IGlyPheylsPheThValylslysaSprysaenglulysSerAepThValGlnle	181
Db	1566	TGCGCTGAAGACAGAGGCCAAAGAAAGACAGGCCCGCAGTACAGACAGCTTCCAA-----	1620
Oy	181	uThrValylslysaSpglugllyaglualaservalGlyalaglyAepHlSGl	201
Db	1621	-----CAGCAAGAACAGGCGCTCCAGGGCGCTTC-----CGCCACCA	1655
Oy	201	ngluproserValGlnurhrlaValGlygluseralaserlulysgluluserglululysgl	221
Db	1656	G-----GTGGAGAGCTTAAGCAGTACAGTACCTGAAGCAAGAAAGAGACAGAGTGAAGA	1706
Oy	221	nserThglulysglnglulgllyThrlulysglnglulnglulnserSerThglulleProle	241
Db	1707	GGTACCGGAGAGAGAGAGGCTCTTAAGGAGAGCGAGATGCGGCTCTCAAGCAGCTGGA	1760
Oy	241	uglnalagluseraPglmlaala-----glulnglualalulysaSpglulylglul	259
Db	1761	TGCTGCAGAGAGAGGAGAGAGGCTCTTAAGGAGAGCGAGATGCGGCTCTCAAGCAGCTGGA	1820
Oy	259	ulysglnglulysglulProThrlulysSerProglu-----SerProSerSerProvalas	277
Db	1821	GGCACTGGAGAGAGAGAGAGGCTGCGCAAGCTGAGATTCGACAGCAACATTCAGGTGGC	1880
Oy	277	nserGlulThrlhSerSerPheylslysaPheThrlhISglYTrpAlaglyTrpArgly	297
Db	1861	TAAATGAACCCCGGAGCAGTGCCTCCAGACTCAATGACACAG-----GCCACGGGGA	1931
Oy	297	slyeThSerPheylslySerlysglulAep-----AspleglulThrlalaglyly	314
Db	1932	GAAAGCAGAGCTGAAGCCGGAAGGAGAGAACTCCAGGCGTGGTGAACAGCCCGCA	1991
Oy	314	sarglysglulnglulalaglyulysValaSpglulnglul-----LysglululysTh	331
Db	1992	GAACAGCATGAGGCCCGAGGCCAGGTTCGAGAGCTTCAGACTCGGCTGTCACCA	2051
Oy	331	rgluproalaserglulnglulnglulProalagluAepThraSpGlnalAargyleuSerAl	351
Db	2052	GCAAAAGCAACTGAAAGAAAGAGGGTGGCCACAGAAAGACAG-----CTCCAGGA	2105
Oy	351	aasPlyrglulysValglulueProleuGlulAepGlnalGlyAspleglulualaserSe	371
Db	2106	GCAGCTCCAGGCCCTCAAAAGAGTCTTGAAAGTCAACCAAGGGGAGGCTTGAA-----	2157
Oy	371	rglulglulysCysAlaProleualathrglulValPheAepGlyulysMetGlnalHhISgl	391
Db	2158	-GAGAGAGAG-----GCGAGGGCTGCAGATGCCCTGGAAAGACAGACAGCTTGATATCT	2210
Oy	391	nglulValValalagluVal-----HisalaserThValGlnlyseTh	405
Db	2211	TGAGCTGAAGCAAGAGACCGAAGCTGTGTGAGCAGCACTAAAGCGGAAAGAAAGAGACT	2270
Oy	405	rglulnglulnglulngllygllyglulalaglyglulysValalValGlnlglThrgl	425
Db	2271	GGAGAGAAAGAGGGCTGGGCCAAGGGGCTGCAGGCTCATTACTCAAGCTTGGGGAGGC	2330
Oy	425	ylulserleuProProglu-----LysleualagluProGlnGlnlValProgl	441
Db	2331	CCATCAGGCTAGACTGAAGTCAAGTCTGCGCGGAGAGTGCAGAGGCCATAGGCTGCCACCA	2390
Oy	441	nglulalaglu---ProalagluglulueMetulysSerAarglulMetCysValserGly	460
Db	2391	CACAGCTAGAGTGAAGTGTGAGCAGCTGTCTCAAGAAAGATGACTGCTGGCGTGCAGGGGA	2450
Oy	460	YasphlsthrluInleuthrlaPleuSerProgluglulysThrluProlyshlAepProgl	480
Db	2451	TGAGGATAGCCAG-----CAAGAGAGAGC-ACAGTATGGCGGCATGTTCC	2494
Oy	480	uglylIlealyserglulValglulMetleuSerSerGlnlulnarglylIleValGlnlyse	500

Db 2495 AGGAACAGCTGATGACTTTGAA----- 2516
 QY 500 rProLeuLysLysLeuPheSerSerSerSerCylLeuLysLysLeuSerCylLysLysGlnLys 520
 Db 2517 -----GGAGGAATGTGTAGA 2530
 QY 520 EGIYLYSAARGLYGLYGLYGLYVAPSLGLUGLPROGLYGLUTRYGLINHSITLHSTHRI 540
 Db 2531 AGGGCCCCCAGGAGCTGCAGAGAGGCCAAGAGAGAGTGGCCAGCATGAAATCCCAAGCG 2590
 QY 540 userProGluSerAlaSPGLUGLILYSGLYGLUSERSERALASERPROGLUGLUPR 560
 Db 2531 AGCTTCA-----GATTAACCCGACAGACAAACTACCTGACCTCAATGCCAAC 2641
 QY 560 OGGLGLUTHRThrCysLeuGlnLysGLYProLeuGlnLysLaProGlnASPGLYGLUALA-- 579
 Db 2642 TGGCCAGAGACT-----CCAGCAGGTCTCAAGAGAGAGAGTCAAGGAGCCCAAGACTTG 2695
 QY 580 -----GlnGlnGlnLysThrThrSerASPGLYGLULysLysAla 591
 Db 2696 CAGATGACCTCTCCACTCTGCAGAGAAAGATGGCTGCCCAACAGCAAGAGGTGGCCGCT 2755
 QY 591 RGLIUGLILYIETHrProTPAlaSerPheLysMetValThrProLysLysArgVala 611
 Db 2756 TGGAGACCTTGTCGCCAAGGACGGGTGAGACACAGAAACAGCTCCCGGAGATGATCA 2815
 QY 611 rGATGProSerGluSer--AspLysGlnGlnLys--LeuGlnLysValLysSerLar 629
 Db 2816 AGGAGCCTGGAGAGGCGAGAGACAGACAGCCCGAGTGGCTGGAAAGACAAACAGGACGCC 2875
 QY 629 hrLeuSerSerThrAspSerThrValSerGlnMetGlnASPGLYValLysThryValGly 649
 Db 2876 AGTTCTGCACACACAGGACGAGCGCTGCAGACTTATGAGCCGGAGCAGAGCATGTGGCA 2935
 QY 649 LUGLILY-----GlnLysProGlnGlnLys 655
 Db 2936 ATGAGCTGAGAACGGCTGCGGAGCCGCGCTGATGAGAGACCCAGGGCGACAGCAGAGGAGC 2995
 QY 655 ----- 655
 Db 2996 GTGGGACAGAGAAAGGAGAGTGTGGCGCGGCTGACCCAGAGAGCGGGAGCCGTGCCAGGCTG 3055
 QY 656 -----P 656
 Db 3056 ACCTTGCCCTGAGAAAGCGCGCCAGACAGACCTTGAGATGCGGCTGCAGAACGCCCTCA 3115
 QY 656 rOLYSAARGYVALAspThrSerValSerTrpGlnLysLeuLysValGlySerSerL 676
 Db 3116 ACGAGCAGCGTGTGAGTTGGCTACCTTCGAAAGAGGACCTGCTCATGTCCTTGACGAA 3175
 QY 676 YSLYSAARGALAsrGlySAlaSerSerSerASPAspGLUGLYGLYProArgThrLeuGlyG 696
 Db 3176 AGGAA-----GGCAAGAGACAGAGATTT-GGCCAAGCTTCGTGCTGTG 3216
 QY 696 IYAspSer-----HisArgAlaGlnGlnLysAlaSerLysAspLysGlnLysA-----GlyT 712
 Db 3217 GAGGAGAGCCAGATATAAAGAGCTGAGAGAACTTCGGCAAAACCGTGTAGAACATGAAGAA 3276
 QY 712 hrAspAlaValProAla-SerThrGlnGlnLysAspGlnLysAlaGlnGlnLysSerSerPro 731
 Db 3277 CA-----GCTGGCTTAAGAAAGAAAGAGAGCAGCATGCTGTCAGAGGCCAA 3324
 QY 732 GlnProAlaGlySerProSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 751
 Db 3325 TCTAGAGCTGCTGGCAGACAGACAGCCAAAGGCCCAAGCTG---GAAAGCATGCGGGCA 3381
 QY 752 LeuValThrProArgLysLysSerLysSerLysLeuGlnGlnLysAlaGlnLys----- 768
 Db 3382 GAGGTGAGCAAGCTGGAAACAGCAATGCTCAGAAAGCAGACAGAGAGGCTGACAGCCTGAA 3441
 QY 769 -----AspSerSerValGlnGlnLeu 775

Db 3442 CGCAGCCTCGAGGCTGAGCGGCTCCCGGGCTGAGCGGACAGTGTCTCTGGAGACTCTG 3501
Qy 776 SerThrGluLeuGluProSerArgGluGluSerTrpValSerIleLysLysPheIlePro 795
Db 3502 CAGGCGCAGTTAGAGAGAGAGCGCCAGAGCTAGGGCACAGTCAGAGTGCCTTAGCCTCG 3561
Qy 796 GlyArgArgLysLysArgAla-----AspGlyLysGlnGlnGlnAlaThrValGluAsp 813
Db 3562 GCCAACGGAGTGGTGCTCTCCGACCAAGGTACAAAGCACACAGCAAGCGGTGAAGAT 3621
Qy 814 SerGlyProValGluLeu-----AsnGluAspAspProAsnValProAlaVal 829
Db 3622 GAGTGGAGCCCGAGTGGCGCGGGCGGCAAGAGCTGAGAGAAAAATAGCCTCATC 3681
Qy 830 ValProLeu---SerGluTyArgAsnAlaValGluArgGluLysMetGluAlaGlnGlyAsn 848
Db 3682 AGCAGCTTGGAGGAGGAGTGTCCATCTCGAATCGCCAGGTCTCTGGAGAGGAGGG--- 3738
Qy 849 ThrGluLeuProGlnLeuGluGlyAlaValTyValSerGluGluLeuSerLysThrLeu 868
Db 3739 ---GAGAGCAAGGAGTTGAAGCGGCTGGTATGCCCGAGTCAGAGAAAGCCAGAGCTG 3795
Qy 869 ValHisThrValSerValAlaValIleAspGlyThrArgAlaValThrSerValGluGlu 888
Db 3796 GAGGAGAGCTGGCGCTCTGC-----AGGCAG 3822
Qy 889 ArgSerProSerTrpIleSerAlaSerValThrGluProLeuGluHisThrAlaGlyGlu 908
Db 3823 AGACAGCCAGCA----- 3834
Qy 909 AlaMetProValGluGluValThrGluLysAspIleIleAlaGluGluThrProVal 928
Db 3835 ---ACAGTGCAGAGCTGCAGAACGCAGCTCTGCTGCGGAGG--- 3876
Qy 929 LeuThrGlnThrLeuProGluGlyLysAspAlaHisAspAspMetValThrSerGlu--- 947
Db 3877 ---AGTGCAGAGCTCGGGAGGAGCTGAGAAACAGCGGGTGGCTTCAGAAAC 3930
Qy 948 ---ValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeuArg 965
Db 3931 CTGCGGCAGGAGTGACCTCACAGGTGAGCGTGCAGGAGGCTGGGCCAAGAATTGAAG 3990
Qy 966 Thr-----GluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSer 983
Db 3991 GCGTGGCAGGAGAAGTTCTTCCAGAAAGAGCAGCGCCCTCTCCACCTGCAGCTC--- 4044
Qy 984 AlaValSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGlu 1003
Db 4045 -----GAGCACACACAGCACAGGCC 4065
Qy 1004 ValGluSerGlyValLeuAspThrGluGluGluArgGlnThrGlnAlaIleLeuGln 1023
Db 4066 CTGCTGAGTGAGTGTGCTCCAGCTAAGCACCTCTGCCAGCAGCTGCAGGCGGAGGCC 4125
Qy 1024 AlaValAlaAspLysValLysGluGluSerGlnValProAlaThrGlnThrValGlnArg 1043
Db 4126 GCTGCCAGAGAAACGCCACCTGAGGAGCTGGAG-----CAGAGCAAGCAGGCC 4173
Qy 1044 ThrGlySerLysAlaLeuGlu----- 1050
Db 4174 GCTGGGGGAGCTGGGGCAGAGCTGCTCGGGGCCAGCGGAGGCTTGGGGAGCTGATTCTCT 4233
Qy 1051 -----LysValGluGluValGluGluAspSerGluValLeuAlaSerGluLys--- 1066
Db 4234 CTGCGGCAGAGAGTGGCAGAGCAGGAGCGAAGACAGCTCAGCAGCTGCGGGCAGAGAGGCC 4293
Qy 1066 ----- 1066
Db 4294 AGCTATGACGAGCAGCTGAGCATGCTGAAGAGGCGCATGCGCTGTGGCAGAGAGAAC 4353
Qy 1066 ----- 1066
Db 4354 CGGGGGCTGGGTGAGCGGGCCAACTTTGGCGGCGAGTTTCTGGAAAGTGGAGTTGGACCAG 4413

Qy 1067 -----GluLysAspValMetProLysGlyProValGlnGluAlaGlyAlaGluHisLeu 1084
Db 4414 GCCCGGAAAGATGATGCTCAAGAGTTGGCAGCGCTACGTCTGATGCTGAGACCCGCTG 4473
Qy 1085 AlaGlnGlySerGluThrGlyGlnAlaThrProGluSerLeuGluValProGluValThr 1104
Db 4474 GCTGAGGTGCGAGGAGACAGACAGACACTGCCCGGAGCTGGAGGTG-----ATGACT 4527
Qy 1105 AlaAspValAspHisValAlaThrCysGlnValIleLysLeuGlnGlnLeuMetGluGln 1124
Db 4528 GCCAAGTATGAGGTGCC-----AAGGTCAAGTCTCTGGAGGAGG 4569
Qy 1125 AlaValAlaProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrPro 1144
Db 4570 CAGCGGTTCAGGAAGAGAGGAGCAAACTCACTGCCAGGTGGAAGAACTGAGTAAGAAA 4629
Qy 1145 LeuAlaAspSerAspThrAlaAspGlyThrGlnGlnAsp-----GluThrIleAspSer 1162
Db 4630 CTGGTGACTCTCACCAGCCCAAGGTGCAGCAGCAAGCTGAAGGTGTCCAGGCT 4689
Qy 1163 Gln-----AspSerLysAlaThrAlaAlaValArgGlnSerGlnValThrGluGluGlu 1180
Db 4690 CAGGAGGCGAGAGCCAGCAGGAGGCCCGCGCTTCAGGCCCGAGTGAATGAATGCA 4749
Qy 1181 AlaAlaThrAlaGlnLysGluGluProSer-----ThrLeu 1192
Db 4750 GCCCAGTTGAGCCAGAAGGAGCAGGCGAGCTGAGCACTATAAGCTGCAGATGGAGAAAGCC 4809
Qy 1193 ProAsnAsnValProAlaGlnGluGluHisGlyGluGlu-----ProGly 1207
Db 4810 AAACACATTTATGATCCCAAGAGCAGCAGACCAAGAGCTGCAGGAGCAGCTGCGGAGC 4869
Qy 1208 ArgAspValLeuGluProThrGlnGlnGluLeuThrAlaAlaAlaValProValLeuAla 1227
Db 4870 CTGAGCAGCTGCAGAGGAAACAAGAGCTCGAGCTGAAGCT----- 4914
Qy 1228 LysThrGluValGlyGlnGluGlyGluValAspTrpLeuAspGlyGluLysValLysGlu 1247
Db 4915 ---GAACGGCTGGGCCATGAGCTACAG-----CAGGCTGGCTGAAGACCAAGGAG 4962
Qy 1248 GluGlnGluValPheValHis-----SerGlyProAsnSerGlnLysAlaAlaAspVal 1265
Db 4963 GCTGAACAGACCTGCGGCCACCTTACTGCCAGGTGCGCAGCTGGAGGACAGGTTGCC 5022
Qy 1266 ThrTyAspSerGluValMetGlyValAlaGlyCysGln-----GluLysGlu 1281
Db 5023 CATGCAACAGCAGCAGCTTCGAGACCTGGGCAAAATTCAGGTGCAACTGATGCTTTAAAG 5082
Qy 1282 SerThrGluValGlnSer-----LeuSer 1289
Db 5083 AGCGTGAGCCCGCAGGCTAAGCCCGAGCTGGAGCTTGAGTATTGACAGCCTGGATCTGAGC 5142
Qy 1290 LeuGluGluGlyGluMetGluThrAspValGluLysGluLysArgGluThrLysProGlu 1309
Db 5143 TGCAGAGGAGGAGCCCACTCAGCTATCACAGCAAGCTGCTCGT---ACCACGCCAGAC 5199
Qy 1310 GlnValSerGluGlyGluGlnGlu-----ThrAlaAlaProGluHisGluGly 1326
Db 5200 GGACACAGCTCCTCGAGAACCCAGCTTACCTATCTCCCA-GCGCTGCCCCCAAGGT 5258
Qy 1327 ThrTyGlyLysProValLeuThrLeuAspMetProSerSerGluArgGlyLysAlaLeu 1346
Db 5259 AGAATCCCTGGAGAGTCTCTACTTACTCCCATCCCTGCTCGGAGTCAAGGCCCCCT--- 5315
Qy 1347 GlySerLeuGlyGlySerProSerLeuPro----- 1356
Db 5316 -----GGAGAGCAGCCTGCACTCCCTGGGAGAGCTTCTCTGAGTCTGGGTCTGTA 5366
Qy 1357 -----AspGlnAspLys 1360
Db 5367 GACCCCGCTCGCTCGTGGCGCACCCAGCAGATCATCAACATCACCATGACCAAGAA--- 5423

QY 221 nSerThrGluLysGlnGluGlyThrLeuLysGlnGlnSerSerThrThrGluIleProLe 241
DB 1707 GTAGCGGAGACAGAGGCACT-----AGCGAGGACCATCGCCAGCAACTGGCCAC 1760
QY 241 uGlnAlaGluSerAspGlnAlaAala-----GluGluGluAlaLysAspGluGluG 259
DB 1761 TGTGCGAGGAGCGAGAGGCTCTTAAGGGAGCGGATGCGGCTCTCAAGCAGCTGGA 1820
QY 259 uLysGlnGluLysGluProThrLysSerProGlu-----SerProSerSerProValAs 277
DB 1821 GGCACCTGGAGAGGAGAGGCTGCCAAGCTGGAGATTCTGCAGCAGCAACTTCAGGTGGC 1880
QY 277 nSerGluThrThrSerSerPheLysLysPhePheThrHisGlyTrpAlaGlyTrpArgLy 297
DB 1881 TAATGAAGCCCGGACAGTCCCGACACCTCAGTGACAG-----GCCAGCGGGA 1931
QY 297 sLysThrSerPheLysLysSerLysGluAsp-----AspLeuGluThrAlaGluLy 314
DB 1932 GAAGGCAGAGCTGAGCCGGAAGGTGGAGAACTCCAGGCTGTGTTGAGACAGCCGCCA 1991
QY 314 sArgLysGluGlnGluAlaGluLysValAspGluGluGlu-----LysGluLysTh 331
DB 1992 GGAACAGCATGAGCCCGAGCCAGGCTGCAGAGCTAGAGTTGCAGCTGCGGCTGAGCA 2051
QY 331 rGluProAlaSerGluGluGlnGluProAlaGluAspThrAspGlnAlaArgLeuSerAl 351
DB 2052 GCAAAAGCAACTGAGAAAGGAGGTGGCCCGAGGAGGACAG-----CTCCAGGA 2105
QY 351 aAspTyrGluLysValGluLeuProLeuGluAspGlnValGlyAspLeuGluAlaSerSe 371
DB 2106 GCAGCTCCAGGCCCTCAAGAGTCTTGAAGTCCACCAAGGCGAGCCTTGA----- 2157
QY 371 rGluGluLysCysAlaProLeuAlaThrGluValPheAspGluLysMetGluAlaHisG 391
DB 2158 -GAGGAGAAG-----CGCAGGCTGCAGATGCCCTGGAAGAGCAGCAGCGTGTATCTC 2210
QY 391 nGluValValAlaGluVal-----HisValSerThrValGluLysTh 405
DB 2211 TGAGCTGAAGCGAGAGACCCGAAAGCTGTGGAGCAGCATAAAGCGGAGCAAGGAGCT 2270
QY 405 rGluGluGluGlnGlyGlyGlyGluAlaGluGlyGlyValValValGluGlyThrG 425
DB 2271 GGAAGAGAGAGGCTGGGCGCAAGGGCTGGAGGCTCGATTACTGACAGCTTGGGAGGC 2330
QY 425 yGluSerLeuProProGlu-----LysLeuAlaGluProGlnGluValProG 441
DB 2331 CCATCAGGCTGAGACTGAAGTCTCGCGCGGAGCTGGCAGAGGCCATGCTGCCCCAGCA 2390
QY 441 nGluAlaGlu-----ProAlaGluGluLeuMetLysSerArgGluMetCysValSerGly 460
DB 2391 CACAGCTGAGAGTGAGTGTGAGCAGCTCGTCAAGAAGTAGCTGCTGGCGGTGACGGTA 2450
QY 460 yAspHisThrGlnLeuThrAspLeuSerProGluGluLysThrLeuProLysHisProG 480
DB 2451 TGAGGTAGCCAG-----CAAGAGGAGGC-ACAGTATGGCGCCATGTTC 2494
QY 480 uGlyIleValSerGluValGluMetLeuSerSerGlnGluArgIleLysValGlnGlySe 500
DB 2495 AGCAACAGCTGATGACTTTGAA----- 2516
QY 500 rProLeuLysLysLeuPheSerSerSerGlyLeuLysLysLeuSerGlyLysLysGlnLy 520
DB 2517 -----GGAGGAATGTGAGA 2530
QY 520 sGlyLysArgGlyGlyGlyAspGluGluProGlyGluTyrGlnHisIleHisThrG 540
DB 2531 AGCCCGCCAGGAGCTGCAGAGGCAAGAGAGAGGTGGCAGGCATAGATCCACAGCG 2590
QY 540 uSerProGluSerAlaAspGlnLysGlyGlyGluSerSerAlaSerSerProGluGluPr 560
DB 2591 AGCTCCA-----GATAAGCCGCGCAGCAGCAACAACCTAGCTGAGCTCCATGCCCAACC 2641
QY 560 oGluGluThrThrCysLeuGluLysGlyProLeuGluAlaProGlnAspGlyGluAla-- 579

DB 2642 TGCCAGAGCACT-----CCAGCAGGTCCAAGAGAAGGAAGTCAGGGCCCAAGACTTG 2695
QY 580 -----GluGluGly-ThrThrSerAspGlyGluLysLysA 591
DB 2696 CAGATGACCTCTCCACTCTGCAGGAAAAGATGGTCCACCAGCAAGAGGTGGCCCGCT 2755
QY 591 rgGluGlyIleThrProTrpTrpAlaSerPheLysLysMetValThrProLysLysArgValA 611
DB 2756 TGGAGACCTTGGTCCGCAAGCAGGTGAGCAGCAGGAAAACAGCCTCCCGGAGTTAGTCA 2815
QY 611 rgArgProSerGluSer---AspLysGluGluGlu---LeuGluLysValLysSerAla 629
DB 2816 AGAGCTCGAGGCGGAGAGCAGACCCGAGTGGCTTGAAGAGCAACAGGAGCGCC 2875
QY 629 hrLeuSerSerThrAspSerThrValSerGluMetGlnAspGluValLysThrValGlyG 649
DB 2876 AGTTCTGCAGCACACAGCAGCGCGCTGCAGGCTATGGAGCGGAGGAGCAGATGGCA 2935
QY 649 luGlu-----GlnLysProGluGlu- 655
DB 2936 ATGAGCTGGAACGGCTGCGGGCGCGCTGTATGAGAGCCAGGGGCGCAGCAGGAGGAGC 2995
QY 655 ----- 655
DB 2996 GTGGGAGCAGAAAGGAGGTGGCGCGCTGACCCAGGAGGGCGCGTGGCCAGGCTG 3055
QY 656 -----P 656
DB 3056 ACCTTGCCCTGGAAAGCGGCGCAGAGCAGAGCTTCAGATGCGGCTGCAGAACGCCCTCA 3115
QY 656 roLysArgArgValAspThrSerValSerTrpAlaLeuIleCysValGlySerSerL 676
DB 3116 ACAGCAGCGTGTGGAGTTGCTTACCTGCAAGAGGCACTGGCTCATGCCCTCAGCGAAA 3175
QY 676 yLysArgAlaArgLysAlaSerSerSerAspAspGluGlyGlyProArgThrLeuGlyG 696
DB 3176 AGAA-----GGCAAGGACGAGGAGTT-GGCCAAGCTTCGTGCTGTG 3216
QY 696 lyAspSer-----HisArgAlaGluAlaSerLysAspLysGluAla-----GlyT 712
DB 3217 GAGGAGCCAGATAAAGAGCTGGAGGAACCTTCGCAACCGTGAAGCAACTGAAGGAA 3276
QY 712 hrAspAlaValProAla-SerThrGlnGlnGlnAspGlnAlaGlnGlySerSerPro 731
DB 3277 CA-----GCTGGCTAAGAAAGAAAAGGAGCAGCATCTGGCTCAGGAGCCCAA 3324
QY 732 GluProAlaGlySerProSerGluGlyGlyGlyValSerThrTrpGluSerPheLysArg 751
DB 3325 TCTGAGGCTGTGGCAGGACAGAGCCCAAGGCCCCCAAGCTG---GAAGCACTGCGGGCA 3381
QY 752 LeuValThrProArgLysLysSerLysSerLysLeuGluGluLysAlaGlu----- 768
DB 3382 GAGGTGAGCAAGCTGGACAGCAATGCCAGAAGCAGCAGGAGGAGCTGACAGCCCTGGAA 3441
QY 769 -----AspSerSerValGluGlnLeu 775
DB 3442 CGCAGCTTCAGGCTGAGCGGGCTCCCGGGCTGAGCGGACAGTGTCTCTGGAGACTCTG 3501
QY 776 SerThrGluIleGluProSerArgGluGluSerTrpValSerIleLysLysPheIlePro 795
DB 3502 CAGGGCCAGTTAGAGGAGAGGCCAGGAGCTAGGCGACAGTCAGAGTGCCTTAGCCTCG 3561
QY 796 GlyArgArgLysLysArgAla-----AspGlyLysGlnGluGlnAlaThrValGluAsp 813
DB 3562 GCCCAAGCGGAGTGGCTGCTCCGCCACCAAGGTATCAAGACCACAGCAAGGCTGAAGAT 3621
QY 814 SerGlyProValIleIle-----AsnGluAspAspProAsnValProAlaVal 829
DB 3622 GAGTGAAGGCCAGGTGGTCCCGGGCGCGGCAAGAGGCTGAGAGGAAAAATAGCCTCATC 3681
QY 830 ValProLeu---SerGluTyrAsnAlaValGluArgGluLysMetGluAlaGlnGlyAsn 848

Db 3682 AGCAGCTTGGAGAGAGGTCTCCATCTGAATCCGACAGTCTCGAGAGAGAGGG--- 3738
 QY ThrGluLeuProGluLeuLeuGluValTyrValSerGluGluLeuSerLysThrLeu 868
 Db 3739 ---GAGAGCAAGAGATTGAAGGGGCTGTGATGGCCGAGTACAGAGAGAGAGAGAGAGCTG 3795
 QY ValHisThrValSerValAlaValIleAspGlyThrArgAlaValThrSerValGluGlu 888
 Db 3796 GAGGAGAGCTGGCGCTGC-----AGGCAG 3822
 QY ArgSerProSerTrpIleSerAlaSerValThrGluProLeuGluHisThrAlaGlyGlu 908
 Db 3823 AGACAGCCAGCA----- 3834
 QY 909 AlaMetProProValGluGluValThrGluLysAspIleIleAlaGluGluThrProVal 928
 Db 3835 -----ACAGTGCAGAGCTGCAGAAAGCAGAGCTGTCTGTGGCGGAGG----- 3876
 QY 929 LeuThrGluThrLeuProGluGlyLysAspAlaHisAspAspMetValThrSerGlu--- 947
 Db 3877 -----AGTGCAGAGCTCCCGGAGGAGGCTGCAGAAACAGCGGCTGGCTTCAGAGAAC 9330
 QY 948 -----ValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeuArg 965
 Db 3931 CTGGCGCAGAGAGCTGACTTCAAGGCTGAGCGTGCAGAGAGCTGGGCGCAAGATTGAAG 3990
 QY 966 Thr-----GluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSer 983
 Db 3991 GGGTGCAGAGAGAAATTCTCCAGAAAGAGAGGCGCTCTCCAGCTGCAGAGCTC----- 4044
 QY 984 AlaValSerGluLeuThrAspSerProAspThrThrGluGluAlaThrProValGluGlu 1003
 Db 4045 -----GAGCACCCGAGCACAGAGGCC 4065
 QY 1004 ValGluSerGlyValLeuAspThrGluGluGluArgGluThrAlaIleLeuGlu 1023
 Db 4066 CTGGTAGAGTAGCTGCTGCAGCTTAAGCACCTTGCAGAGAGCTGCAGAGCCAGAGAGGCC 4125
 QY 1024 AlaValAlaAspLysValLysGluGluSerGluValProAlaThrGluThrValGluArg 1043
 Db 4126 GCTGCGCAGAGAAAGCCAGCCGTGAGAGCTGGAG-----CAGAGCAAGAGGCC 4173
 QY 1044 ThrGlySerLysAlaLeuGlu----- 1050
 Db 4174 GCTGGGGGAGCTGGGGCAGAGCTGCAGGCGCCAGCGGAGCTTGGAGAGCTGATTCT 4233
 QY 1051 -----LysValGluGluValGluGluAspSerGluValLeuAlaSerGluLys--- 1066
 Db 4234 CTGGCGCAGAGAGTGGCAGAGCAGAGCAAGCTCAGCAGCTGGCGGCGAGAAAGGCC 4293
 QY 1066 ----- 1066
 Db 4294 AGCTATGCAGAGCAGCTGAGCATGTGAAGAGGCGCATGCTGTGGCAGAGAGAAC 4353
 QY 1066 ----- 1066
 Db 4354 CGGGGCTGGGTGAGCGGCGCACTTGGCGGCACTTCTGGAAGTGAAGTGAACAG 4413
 QY 1067 -----GluLysAspValMetProLysGlyProValGluGluAlaGluHisLeu 1084
 Db 4414 GCCCGGGAAGATGATCTCAAGAGTTGGCAGCCGTACGTCTATGCTGTGAGACCCGCTG 4473
 QY 1085 AlaGluGlySerGluThrGluAlaThrProGluSerLeuGluValProGluValThr 1104
 Db 4474 GCTGAGGTGAGGAGAGAGCAGAGACAGCTGCCGGAGCTGGAGGTG-----ATGACT 4527
 QY 1105 AlaAspValAspHisValAlaThrCysGluValIleLysLeuGluGluLeuMetGluGlu 1124
 Db 4528 GCCAAGTATGAGGTGCC-----AAGGTCAAAGTCTCGAGAGAGAG 4569
 QY 1125 AlaValAlaProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrPro 1144
 Db 4570 CAGCGGTTCAGGAGAGAGAGCAAGAACTCAGCTGCCCCAGGTGAGAGAACTGAAGAGAA 4629

QY 1145 LeuAlaAspSerAspThrAlaAspGlyThrGluGluAsp-----GluThrIleAspSer 1162
 Db 4630 CTGGCTGACTGTGACCAAGCCAGAGAGGTGCAGAGAGAGAGAGCTGCAGAGCT 4689
 QY 1163 Gln-----AspSerLysAlaThrAlaAlaValArgGlnSerGluValThrGluGluGlu 1180
 Db 4690 CAGGAGGCGAGAGCCAGAGAGAGGCGCTTCCAGGCGCCAGCTGAATGAAGTGA 4749
 QY 1181 AlaAlaThrAlaGluLysGluGluProSer-----ThrLeu 1192
 Db 4750 GCCCAGTTGAGCAGAAAGAGAGAGAGAGAGCTGAGCACTTAAGCTGAGATGAGAGAAAGCC 4809
 QY 1193 ProAsnAsnValProAlaGluGluGluHisGlyGluGlu-----ProGly 1207
 Db 4810 AAACACATTATGATTCAGAAAGAGAGAGAGAGAGCTGCAGAGAGAGAGAGAGC 4869
 QY 1208 ArgAspValLeuGluProThrGluGluGluThrAlaAlaValProValLeuAla 1227
 Db 4870 CTGGAGCAGCTGCAGAGAGAGAGAGAGAGAGAGCTGCAGAGCTGAAGCT----- 4914
 QY 1228 LysThrGluValGlyGluGluGlyGluValAspTrpLeuAspGlyGluLysValLysGlu 1247
 Db 4915 ---GAACGCTGGGCGCATGAGCTAAG-----CAGGCTGGGCTGAAGCAGAGAG 4962
 QY 1248 GluGluGluValPheValHis-----SerGlyProAsnSerGluLysAlaAlaAspVal 1265
 Db 4963 GCTGAACAGAGCTTGGCGGACCTTACTGTCAGAGTGCAGAGCTTGGAGCAGAGTTGCC 5022
 QY 1266 ThrTrpAspSerGluValMetGlyValAlaGlyCysGlu-----GluLysGlu 1281
 Db 5023 CATGCAGACACAGAGCTTGCAGAGCTGGGCAATTCCAGGTGGCACTGATGCTTTAAG 5082
 QY 1282 SerThrGluValGlnSer-----LeuSer 1289
 Db 5083 AGCGTGAAGCCCAAGCTTAAGCCAGGTGAGCTTGAAGATTGACAGCTGAGCTGAGC 5142
 QY 1290 LeuGluGluGlyGluMetGluThrAspValGluLysGlyLysGluThrLysProGlu 1309
 Db 5143 TCCGAGAGGGAGGCCCACTCACTGATCAACAGCAAGCTGCCCTCGT---ACCAGCCAGAAC 5199
 QY 1310 GlnValSerGluGluGluGluGluGlu-----ThrAlaProGluHisGluGly 1326
 Db 5200 GGCACCAAGGCTCCGAG 5258
 QY 1327 ThrTrpGlyLysProValLeuThrLeuAspMetProSerSerGluArgGlyLysAlaLeu 1346
 Db 5259 AGAATCCCTGAGAGATCTTACTTCACTCCATCCCATCCCTGCTGGAGTCAAGGCCCT-- 5315
 QY 1347 GlySerLeuGlyLysProSerLeuPro----- 1356
 Db 5316 -----GGAGAGCAGAGCTGAGACTCCCTGGAGAGAGCTTCTGAGAGCTGGGTGA 5366
 QY 1357 -----AspGlnAspLys 1360
 Db 5367 GACCGGCTCCGCTCGTGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5423
 QY 1361 AlaGlyCysIleGluValGlnValGlnSerLeu-AspThrThrValThrGluThrAlaGlu 1380
 Db 5424 GCTAGATGTGAAG 5483
 QY 1380 ValAlaGluLysValIle-----GluThrValValIleSerGluThrGlyGlu 1396
 Db 5484 TCTTCCCAAGGTGAGCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5541
 QY 1396 uSerProGluCysValGlyAlaHisLeuLeu----- 1406
 Db 5542 -TCTCCGATTTATGGCACTAGCCCTGCTCAGCTTCCCTGAGTACCGGCCCAAGAGAGAG 5600
 QY 1407 ---ProAlaGluLysSerSerAlaThrGlyGlyHisTrpThrLeuGluHisAlaGluLys 1425
 Db 5601 CAGTTCTGCTCGTCTCCAG 5659

QY 1425 pThrValPro-----LeuGlyProGluSe 1433
Db 5660 GCTTCTACATGGGACCTTGGCAGATGAGCTGAGCTGGATGAGTGAACCGCATTTG 5719
QY 1433 rGlnAlaGluSerIleProIleValThrProAlaProGluSerThrLeuHisProAs 1453
Db 5720 CAGAGCTGCAGCAGCAGCAATCGAGTGTGCCCCACACATCTGAAGACCTGTATCCCTGG 5779
QY 1453 pLeuGlnGlyGluSerAlaSerGlnArgGluArg---SerGluGluGluAspLysPr 1472
Db 5780 AGTCCAGGCTTCCCTGAGCTGGGCCACCATCAGATGAGGAGATGAACACTGGAGACC 5839
QY 1472 oAsp-----AlaGlyProAspAlaAspGly 1480
Db 5840 CCAAGAGACCTGCGCGGAGCCAGCATGCGAGCCATCCAGATAGCGGAGGG 5891

RESULT 9
US-08-467-781-3
; Sequence 3, Application US/08467781
; Patent No. 5780596
; GENERAL INFORMATION:
; APPLICANT: TOUTKATLY, GARY
; APPLICANT: LIDGARD, GRAHAM P
; TITLE OF INVENTION: NOVEL MALIGNANT CELL TYPE MARKERS OF THE
; TITLE OF INVENTION: INTERIOR NUCLEAR MATRIX
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT
; STREET: 125 HIGH STREET
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,781
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: PITCHER ESQ, EDMUND R
; REGISTRATION NUMBER: 27,829
; REFERENCE/DOCKET NUMBER: MTP-013
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6306 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..6306
; PUBLICATION INFORMATION:
; AUTHORS: COMPTON, DUANE A
; AUTHORS: SZILAK, ILYA
; AUTHORS: CLEVELAND, DON W
; TITLE: PRIMARY STRUCTURE OF NUMA, AN INTRANUCLEAR
; TITLE: PROTEIN THAT DEFINES A NOVEL PATHWAY FOR
; TITLE: SEGREGATION OF PROTEINS AT MITOSIS
; JOURNAL: J. Cell Biol.
; VOLUME: 116
; PAGES: 1395-1408
; DATE: 1992
US-08-467-781-3

Alignment Scores:

Pred. No.: 1.98e-19 Length: 6306
Score: 415.00 Matches: 373
Percent Similarity: 35.28% Conservative: 255
Best Local Similarity: 20.96% Mismatches: 663
Query Match: 5.14% Indels: 493
DB: 1 Gaps: 72

US-09-902-432-4 (1-1596) x US-08-467-781-3 (1-6306)

QY 28 GlyHisGlyProAlaAlaGluAlaSerGlyAlaAlaGlyAspProAlaAspPro 47
Db 1050 GGCCACTCAGGAGTGGCTAGAGAAGCAGGCCAGCTGGAGAGGAGCTCAGCCGAGCCCT 1109
QY 48 Ala-ThrLys-----LeuProGlnLysAsn-----GlyGlnLeuSerSerVa 61
Db 1110 GCAGCAAGAATAATGCTTTGAAGAGAAGAACGAAATCTTTCAGGGAAAACTTTTCACAGCT 1169
QY 61 1-----AsnGlyValAlaGluGlnGlyAspVal----- 70
Db 1170 GGAAGAACAATTTGTCACAGCTGTCAGGATTAACCCACCCAGGAGAGGGGAGGTGCTGGG 1229
QY 71 -----HisValGlnGluGluAsnGlnGluGly----- 79
Db 1230 TGATGCTTTCAGCTGGAACCTTGAAGCAAGAGCGGAGCCACTCTTCTGCTCAAAACAACAC 1289
QY 80 -----GlnGluGluGluValValAspGluAspValGlyGlnArgGluSerGluAs 96
Db 1290 ACAGCTCCAAAGCCAGGCTAGAGATGCTGGAGACTGAGCAGGAGCCAGCAGGAAGCCAAGCT 1349
QY 96 pValArgGluLysAspArgValGluGlu----- 105
Db 1350 GCTTGTGTCGCGGGCCACTTCCAAAGAAAAGCAGCAGCTGTCTAGCCTGTATCACTGA 1409
QY 106 -----MetAlaAlaAsnSerThrAlaValAlaGluAspIleThrLysAspGlyG 121
Db 1410 CCTGCAGAGCTCCATCTCAACCTCAGCCAGCCAGGAGAGAGCTGGAGCGCCTCCCA 1469
QY 121 nGluGluThrSerGluIleGluGlnIleProAlaSerGluAsnValGluGluMe 141
Db 1470 GGCTCATGGGCGCGTTCAGCTGCCAGGTGGCTCTCTGACCTCTGAGCTCACCACACT 1529
QY 141 tValGlnProAlaGluSerGlnAlaAsnAspValGlyPheLysLysValPheLysPheVa 161
Db 1530 CAATGCCACCATCCAGCAACAGGATCAAGAACTG-----GC 1565
QY 161 lGlyPheLysPheThrValLysLysAspLysAsnGluLysSerAspThrValGlnLeuLe 181
Db 1566 TGGCCTGAAGCAGCAGCGCCAAAGAGAGAGCAGGCCAGCTAGCAGACACCTCCAA----- 1620
QY 181 uThrValLysLysAspGluGluGlyAlaGluAlaSerValGlyAlaGlyAspHisG 201
Db 1621 -----CAGCAAGAACAGCGCTCCAGGGGCTC---CGCCACCA 1655
QY 201 nGluProSerValGluThrAlaValGlyGluSerAlaSerLysGluSerGluLeuLysG 221
Db 1656 G-----GTGGAGCAGCTAAGCAGTAGCTAGCCTGAAGCAGAGGAGCAGCAGTTGAAGGA 1706
QY 221 nSerThrGluLysGlnGluGlyThrLeuLysGlnGluGlnSerSerThrGluIleProLe 241
Db 1707 GTTAGCGGAGAGCAGGAGGCACT-----AGGCGAGGACCATGCCAGCACTGGCCAC 1760
QY 241 uGlnAlaGluSerAspGlnAlaAla-----GluGluGluAlaLysAspGluGluG 259
Db 1761 TGCTGCAGAGGAGCAGAGAGCGCTCTTAAGGGAGCGGGATGCGGCTCTCAAGCAGCTGGA 1820
QY 259 uLysGlnGluLysGluProThrLysSerProGlu-----SerProSerSerProValAs 277
Db 1821 GGCACTGGAGAAGGAGAGGCTGCCAAGCTGGAGATTCTGCAGCAGCACTTCAGGTGCGC 1880
QY 277 nSerGluThrThrSerSerPheLysLysPhePheThrHisGlyTrpAlaGlyTrpArgly 297
Db 1881 TAATGAAGCCCGGACAGTGGCCAGACCTCACTGAGCACAG-----GCCACGCGGA 1931


```

OY 227 slysthrSerPheIlySlySerIySgluasp-----AspleuGluThrAlaGluIy 314
Db 1932 GAAGCAGAGCTGAAGCCGGAAGGTGAGAACTCCAGGCTGTGTGAGACGCGCGCA 1991
OY 314 sArgIySgluIngluAlaGluIyValAspIgluInglu-----LysGluIyThr 331
Db 1992 GGAACGAGCTAGAGCCCGAGGCCCGGCTTCAGAGCTAGAGCTTCGAGCTCGGTGAGA 2051
OY 331 rGluProIaSerIyGluIngluIngluProIaGluAspThrAspGlnAlaArgIyLeuSerAl 351
Db 2052 GCAAAAGCACTGAGAAAGAAAGAGGTGGCCGAGAGAACAGCAG-----CTCCAGGA 2105
OY 351 aAspIyrgIuIySValGluIyLeuProIeuIyAspGlnValGlyAspLeuGluAlaSerSe 371
Db 2106 GCAGCTCCAGGCCCTCCAAAGAGTCTTGAAGTCCACCAAGGAGGAGCCTTGAA----- 2157
OY 371 rGluIyIyScySaIaProIeuAlaThrGluValPheAspGluIyMetGluAlaHisG1 391
Db 2158 -GAGGAGAAAG-----CGCAGGCTCCAGATGCCCTGGAAGAGCAGCAGCTTGTATCTC 2210
OY 391 nGluValAlaIaGluIyVal-----HisValSerThrValGluIyThr 405
Db 2211 TAGAGTGAAGAGAGAGACCCGAGCCTGTGGAGCAGCATTAAGCGGAGCAAGAGAGCT 2270
OY 405 rGluIyGluIngluIngluIyGlyIyGlyIyGluAlaGluIyGlyValValGluIyThrG1 425
Db 2271 GGAAGAGAGAGAGGCTGGCGCAGAGGGGCTGGAGGCTGATTAATCAGCTTGGGGAGGC 2330
OY 425 yGluSerIyLeuProIyProIy-----LysIyLeuIaGluIyProIyGluIyValProI 441
Db 2331 CCATCAGCTGAGACTGAAGTCTCTCGCGGAGGAGCTGGCAGAGGCCATAGCTGCCAGCA 2390
OY 441 nGluIaIaGluIy-----ProIaGluIyLeuMetIySerArgIyMetCysValSerIyG1 460
Db 2391 CACAGCTGAGAGTGAAGTGAAGCAGCTCCGCAAAAGAAAGTGCCTGGGCTGAGCGGTA 2450
OY 460 yAspHisThrGluIyLeuThrAspLeuSerProIyGluIyIyThrIyProIyHisProI 480
Db 2451 TGAAGTATACAG-----CAAGAGAGGC-ACAGTATGAGCGCATTTTCC 2494
OY 480 uGlyIyIyValSerGluValGluMetLeuSerSerIyGluIyArgIyIyValGluIySe 500
Db 2495 AGAACAAGCTGATGACTTTGA----- 2516
OY 500 rProIeuIySlyIyLeuPheSerSerSerIyLeuIySlyLeuSerGlyIySlyGluIy 520
Db 2517 -----GAGGAATGTGAGA 2530
OY 520 sGlyIySArGlyIyGlyIyGlyIyAspGluIyProIyGlyIyTrGlnHisIyHisThrG1 540
Db 2531 AGGCCCCGAGGAGCTGACAGAGCAAGAGAGAGTGGCAGGCAATAGATCCACAGCG 2590
OY 540 uSerProIyLeuSerAlaAspGluIyIySglIyIyLeuSerSerIyAlaSerSerProIyGluIyP 560
Db 2591 AGCTCCA-----GATTAAGCCGCGCAGCAGAACTAGCTGAGCTCCAGCCCAACC 2641
OY 560 oGluIyIyThrIyCysLeuGluIySglIyProIeuIyAlaProIyAspGlyIyGluAla-- 579
Db 2642 TGGCCAGACACT-----CCAGCAGTCCCAAGAGAAAGAAAGTCAAGGCCCAAGCTTG 2695
OY 580 -----GluIyGluIy-ThrThrSerAspGlyIyGluIyLeuIy 591
Db 2696 CAGATGACTCTTCACCTTCAGAGAAAGATGCTCCACACAGCAAAAGAGTGGCCGCT 2755
OY 591 rGluIyIyIyLeuThrProIyAlaSerPheIySlyMetValThrProIySlyArgVala 611
Db 2756 TCGAGACTTGTGCTCCCAAGGAGTGAAGAGAGAAACAGCTCCCGGAACTTATGTA 2815
OY 611 rArgProSerIyLeuSer-----AspIySgluIyGluIy--LeuGluIySValIySerAla 629
Db 2816 AGGAGCTCTGAGAGGAGAGAGACAGACAGCCAGTGGCTGGAAGAGCAACAGGAGCGC 2875
OY 629 hrLeuSerSerThrAspSerThrValSerGluMetGluAspGluIyValIyThrValIyG 649

```

```

Db 2876 AGTTTCAGACACACAGCAGCGCTGACAGCTATGAGCGGAGAGCAGACAGATGGCA 2935
OY 649 IyGlu-----GluIySerProIyGluIy 655
Db 2936 ATGAGCTGGAACCGGCTGCGGCGCGCTGATGAGAGAGCCAGGGGAGAGAGAGGAGC 2995
OY 655 ----- 655
Db 2996 GTGGGACAGAGAAAGGAGGTGGCGGCTGACCCAGAGAGCGGCGCTGCCAGGCTG 3055
OY 656 -----P 656
Db 3056 ACCTTGCCCTGAGAAAGCGCGCCAGAGCAGCTTGAGATGCGCTGAGAAACCCCTCA 3115
OY 656 rOlySArgArgValaAspThrSerValSerIyProIyAlaLeuIyCysValGlySerIy 676
Db 3116 ACAGCAGCGTGTGAGATTGCTACCTTCGCAAGAGGAGCAGCTGGCTCATGCGTCAAGAAA 3175
OY 676 ySlySArgAlaArgIyAlaSerSerSerAspAspGluIyGlyProArgThrIyLeuIyG 696
Db 3176 AGGAA-----GCAAGAGACAGAGAGTT-GGCCAAGCTTCGTGTCTG 3216
OY 696 IyAspSer-----HisArgIaGluIyAlaSerIyAspIySgluAla-----GlyT 712
Db 3217 GAGCAGCCCGAGATTAAGAGCTGAGAGAACTTCGCAACCGTGTAAACCACTGAAGAA 3276
OY 712 hrAspAlaValProIa-SerThrGluIyGluIyAspGlnIaGluIySerSerSerPro 731
Db 3277 CA-----GCTGGCTTAAGAAAGAAAGAGACGATCTGGCTTCAGAGAGCCCA 3324
OY 732 GluProIaGlySerProSerGluIyGluIyValSerThrTrpGluSerPheIySArg 751
Db 3325 TCTGAGGCTGCTGGCAGAGACAGGCAACAGGCCCAAGCTG--GAAGCACTGGCGCA 3381
OY 752 LeuValThrProArgIyIySlySerIySerIySlyLeuGluIyIyValIyGluIy 768
Db 3382 GAGGTAGCAAGCTGGAACAGCAATGCCAAGAGCAGCAGAGAGGCTGACAGCCTGGA 3441
OY 769 -----AspSerSerValGluIyLeu 775
Db 3442 CGCAGCTCGAGGCTGAGCGGCGCTCCGCGGCTGAGCGGAGAGTGCCTGAGAGCTCG 3501
OY 776 SerThrGluIyGluProSerArgIyGluIySerTrpValSerIyIyLeuIySlyPheIyPro 795
Db 3502 CAGGCGCAGTGAAGAGAGAGGCCAGAGCTAGGGCAGTCAAGTGCCTTACCTCG 3561
OY 796 GlYArGArgIySlyArgAla-----AspGlySgluIngluAlaThrValGluIyP 813
Db 3562 GCCCAACGGGAGTGTGCTGCTTCGCGCACCAAGGTACAAAGACACAGAGGCTGAAGT 3621
OY 814 SerGlyProValGluIy-----AsnGluAspAspProAsnValProIaVal 829
Db 3622 GAGTGAAGGCCAGGTGGCCCGGCGCGGCAAGAGGCTGAGAGAGAAATAGCTCATC 3681
OY 830 ValProIeu--SerGluIyTrsnaIaValGluArgIyIyMetGluIyAlaIngluIyAsn 848
Db 3682 AACAGCTTGAAGAGAGAGTGTCCATCTGAATCGCCAGTCTGGAAGAGAGGGG-- 3738
OY 849 ThrGluIyLeuProIyLeuIyGluIyAlaValIyTrValSerGluIyLeuSerIyThrIyLeu 868
Db 3739 ---GAGAGCAAGAGAGTGAAGCGCTGTGATGGCGGATCGAGACAGAACGCAAGCTG 3795
OY 869 ValHisThrValSerValAlaValIyIyAspGlyThrArgAlaValThrSerValGluIy 888
Db 3796 GAGGAGAGTGGCGCTGCTGC-----AGGCA 3822
OY 889 ArgSerProSerTrpIyIySerAlaSerValThrGluProIeuGluHisThrAlaGlyIy 908
Db 3823 AGACAGCGAGCA----- 3834
OY 909 AlaMetProIyValGluIyValThrGluIySAspIyIyIyAlaGluIyIyThrProIyVal 928

```

Db 3835 -----ACAGTGCAGAGCTGCAGAACGCGAGCTCTGCTCTGCGGAGG----- 3876
Qy 929 LeuThrGlnThrLeuProGluGlyLysAspAlaHisAspMetValThrSerGlu--- 947
Db 3877 -----AGGTGCAGAGCTCCGGAGGGAGCTGAGAAACAGCGGGTGGCTTCAGAGAAC 3930
Qy 948 -----ValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeuArg 965
Db 3931 CTGGCGAGGAGTGAACCTCACAGGTGAGCGTCCGAGGAGGCTGGGCCCAAGAAATTGAAG 3990
Qy 966 Thr-----GluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSer 983
Db 3991 GCCTGCAGGAGAGTCTTCAGAAAGACGAGCGCTCTCCACCTCGCAGCTC----- 4044
Qy 984 AlaValSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGlu 1003
Db 4045 -----GAGCACACACAGCACAGGCC 4065
Qy 1004 ValGluSerGlyValLeuAspThrGluGluGluGluArgGlnThrGlnAlaLeuGln 1023
Db 4066 CTGGTGAGTGAGCTGTGCGCAGTAAAGCACTCTGCGCAGAGCTGCGAGCGGCGAGAGGCC 4125
Qy 1024 AlaValAlaAspLysValLysGluGluSerGlnValProAlaThrGlnThrValGlnArg 1043
Db 4126 GCTGCCGAGAAACGCCACCGTGGAGCTGGAG-----CAGACAGACGAGGCC 4173
Qy 1044 ThrGlySerLysAlaLeuGlu----- 1050
Db 4174 GCTGGGGAGCTGCGGGCAGAGCTGCTCGCGGCCAGCGGAGCTTGGGGAGCTGATTCTCT 4233
Qy 1051 -----LysValGluGluValGluAspSerGluValLeuAlaSerGluLys--- 1066
Db 4234 CTCCGCGAAGAGTGGCAGACGAGCGAAGCAAGCTCAGCAGCTGCGGGCAGAGAGGCC 4293
Qy 1066 ----- 1066
Db 4294 AGCTATGCAGACAGCTGACATGCTGAAGAAGCGCATGGCTGTGCGCAGAGGAGAAC 4353
Qy 1066 ----- 1066
Db 4354 CGGGGGCTGGTGAGCGGGCCAACTTGGCGCGGAGTTTCTGGAAAGTGGAGTTGGACCAG 4413
Qy 1067 -----GluLysAspValMetProLysGlyProValGlnGluAlaGlyAlaGluHisLeu 1084
Db 4414 GCCCGGGAAGATGTCTCAAGAGTTGGCAGCGCTACGTCTGATGTGAGACCCGCTG 4473
Qy 1085 AlaGlnGlySerGluThrGlyGlnAlaThrProGluSerLeuGluValProGluValThr 1104
Db 4474 GCTGAGGTGCGCAGAGACAGACAGACCTGCCCCGGGAGCTGGAGGTG-----ATGACT 4527
Qy 1105 AlaAspValAspHisValAlaThrCysGlnValIleLysLeuGlnGlnLeuMetGluGln 1124
Db 4528 GCCAAGTATGAGGTGCC-----AAGTCAAGTCTCTGGAGGAGG 4569
Qy 1125 AlaValAlaProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrPro 1144
Db 4570 CAGCGGTTCCAGGAAGAGAGCGCAAACTCACTGCCAGGTGGAAGAACTGAGTAAGAAA 4629
Qy 1145 LeuAlaAspSerAspThrAlaAspGlyThrGlnGlnAsp-----GluThrIleAspSer 1162
Db 4630 CTGGTGACTCTGACCAAGCCAGCAGGTGCAGCAGCAGCAAGCTGAAGGTGTCCAGGCT 4689
Qy 1163 Gln-----AspSerLysAlaThrAlaAlaValArgGlnSerGlnValThrGluGluGlu 1180
Db 4690 CAGGGAGCGCAGAGCCAGCAGAGGCCCGCAGCGCTTCAGGCCCGCAGTGAATGACTGCAA 4749
Qy 1181 AlaAlaThrAlaGlnLysGluGluProSer-----ThrLeu 1192
Db 4750 GCCCAGTTGAGCCAGAGAGGACGAGCGACTGAGCACTATAAGCTGCAGATGGAGAAAGCC 4809
Qy 1193 ProAsnAsnValProAlaGlnGluGluHisGlyGluGlu-----ProGly 1207
Db 4810 AAAACACATTATGATGCCAAGAGACGACAGAACCAAGAGCTGCAGGAGCAGCTGCGGAGC 4869

Qy 1208 ArgAspValLeuGluProThrGlnGlnGluLeuThrAlaAlaValProValLeuAla 1227
Db 4870 CTGGAGCAGCTGCAGAAAGGAAAAACAAGAGCTCCGAGCTGAAGCT----- 4914
Qy 1228 LysThrGluValGlyGlnGluGlyGluValAspTrpLeuAspGlyGluLysValLysGlu 1247
Db 4915 ---GAACGGCTGGCCATGAGTACAG-----CAGCTGGGTGAGACCAAGGAG 4962
Qy 1248 GluGlnGluValPheValHis-----SerGlyProAsnSerGlnLysAlaAlaAspVal 1265
Db 4963 GCTGAACAGACCTGCCCGCACCTTACTGCCAGGTGCGAGCTGGAGGCACAGGTTGCC 5022
Qy 1266 ThrTyAspSerGluValMetGlyValAlaGlyCysGln-----GluLysGlu 1281
Db 5023 CATGCAGACAGCAGCTTCGAGACCTGGCGCAAAATTCAGGTGGCAACTGATGCTTTAAAG 5082
Qy 1282 SerThrGluValGlnSer-----LeuSer 1289
Db 5083 AGCGTGAGCCCGAGCTAAGCCCCAGCTGGACTTGAGTATTGACAGCCTGGATCTGAGC 5142
Qy 1290 LeuGluGluGlyGluMetGluThrAspValGluLysGluLysArgGluThrLysProGlu 1309
Db 5143 TCGCAGAGGGGAGCCCACTCAGTATCACAGCAAGCTGCCTCGT---ACCAGCCAGAC 5199
Qy 1310 GlnValSerGluGluGlyGluGlnGlu-----ThrAlaAlaProGluHisGluGly 1326
Db 5200 GGCACAGCGCTCCTGGAGAACCCAGCTCACCTATCTCCCA-GCGCTGCCCCCAAGGT 5258
Qy 1327 ThrTyGlyLysProValLeuThrLeuAspMetProSerSerGluArgGlyLysAlaLeu 1346
Db 5259 AGAATCCCTGGAGAGTCTCTACTTCACTCCCATCCCTGCTCGGAGTCAGGCCCCCT--- 5315
Qy 1347 GlySerLeuGlyGlySerProSerLeuPro----- 1356
Db 5316 -----GGAGAGCAGCTGAGCTCCCTGGAGAGCTTCTCTGAGACTCGGCTCGTAA 5366
Qy 1357 -----AspGlnAspLys 1360
Db 5367 GACCCGCTCGCTCGTGGCGCACCGCAGATCATCAACATCACCATGACCAAGAA--- 5423
Qy 1361 AlaGlyCysIleGluValGlnValGlnSerLeu-AspThrThrValThrGlnThrAlaGln 1380
Db 5424 GCTAGATGTGGAAGAGCAGCAGCGCCCAACTCATCTTCTACAGCAGCGGTCTGCTCC 5483
Qy 1380 uAlaValGluLysValIle-----GluThrValIleSerGluThrGlyGln 1396
Db 5484 TGCTTCCAGGTAGCTGCGGAGCCACCTCTCTACTCAGTCTCTAGCTGCGCTGGGT--- 5541
Qy 1396 uSerProGluCysValGlyAlaHisLeuLeu----- 1406
Db 5542 -TCTCCGATTATGGCAACTCAGCCCTGCTCAGCTTGGCTGCTACCGCCGCCACTCG 5600
Qy 1407 ---ProAlaGluLysSerSerAlaThrGlyHisIleThrLeuGlnHisAlaGluAs 1425
Db 5601 CAGTTCTGCTGCTGCTTCCAGCGCGG-GGTGTCCAGTGGGGCCCTCCAGGAAGGAACA 5659
Qy 1425 pThrValPro-----LeuGlyProGluSe 1433
Db 5660 GCTTCTACATGGGCATTCGCCAGGTAGCCCTGAGCAGCTGGATGACTGGAACCGCATTTG 5719
Qy 1433 rGlnAlaGluSerIleProIleIleValThrProAlaProGluSerThrLeuHisProAs 1453
Db 5720 CAGAGCTGCAGCAGCGCAATCGAGTGTGCCCCCACTCATGAAGACCTGCTATCCCTGG 5779
Qy 1453 pLeuGlnGlyGluIleSerAlaSerGlnArgGluArg---SerGluGluGluAspLysPr 1472
Db 5780 AGTCCAGGCTTCTCCCTGAGCTGGGCACCATCACAGATGAGGAGATGAAAACTGGAGACC 5839
Qy 1472 oAsp-----AlaGlyProAspAlaAspGly 1480
Db 5840 CCAAGAGACCTTGGCCCGCAGCAGCATGCAGCAATCCAGATAGCCGAGGG 5891

```

RESULT 10
US-08-483-924-3
Sequence 3, Application US/08483924
Patent No. 5882876
GENERAL INFORMATION:
APPLICANT: TOUKATLY, GARY
APPLICANT: LIDGARD, GRAHAM P
TITLE OF INVENTION: NOVEL MALIGNANT CELL TYPE MARKERS OF THE
TITLE OF INVENTION: INTERIOR NUCLEAR MATRIX
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: TESTA, HURWITZ & THIBEAULT
STREET: 125 HIGH STREET
CITY: BOSTON
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/483,924
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: PITCHER ESQ, EDMUND R
REGISTRATION NUMBER: 27, 829
REFERENCE/DOCKET NUMBER: MTP-013
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 6306 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..6306
PUBLICATION INFORMATION:
AUTHORS: COMPTON, DUANE A
AUTHORS: SZILAK, ILVA
AUTHORS: CLEVELAND, DON W
TITLE: PRIMARY STRUCTURE OF NUMA, AN INTRANUCLEAR
TITLE: PROTEIN THAT DEFINES A NOVEL PATHWAY FOR
TITLE: SEGREGATION OF PROTEINS AT MITOSIS
JOURNAL: J. Cell Biol.
VOLUME: 116
PAGES: 1395-1408
DATE: 1992
US-08-483-924-3

Alignment Scores:
Pred. No.: 1,98e-19 Length: 6306
Score: 415.00 Matches: 373
Percent Similarity: 35.28% Conservative: 255
Best Local Similarity: 20.96% Mismatches: 663
Query Match: 5.14% Indels: 493
Gaps: 72

US-09-902-432-4 (1-1596) x US-08-483-924-3 (1-6306)
QY 28 GYHISGLYProAlaIaGluAlaSerGlyAlaIaGlyAspProAlaIaAspPro 47
Db 1050 GCCCACTCAGAGTGGCTGAGAGAGCAGGCCCTGAGAGAGAGCTCAGCCAGCCCT 1109
QY 48 Ala-ThrIys-----LeuProGlnIysAsn-----GlyGlnIysSerSery 61
Db 1110 GCAGGACAAAGAAATGCTTGAAGAGAGAGAAATCTCTTCAGGAGAAACTTTCACAGCT 1169

```

```

QY 61 1-----AsnGlyValAlaGluGlnGlyAspVal----- 70
Db 1170 GAAAGAACACTTGTCCAGCTGACAGATACCAACCCCAAGAGAGAGGCTGTGGG 1229
QY 71 -----HisValGlnGlnGluIaAsnGlnGluIy----- 79
Db 1230 TGATGCTTGCAGCTGGAACCTTGAACCAAGAGCCACTCTGTGCAACACAC 1289
QY 80 -----GlnGlnGluIaValAlaIaAspGluIaValGlnIaIaGlnIaIa 96
Db 1290 ACAGCTCCAGCAGGAGTGAAGATGCTGAGACTGAGAGGCGCAGAGCAAGCT 1349
QY 96 pValIaIaGluIaIaAspIaIaGluIaIa----- 105
Db 1350 GCTTGTCTAGCGGGCCACTTCTGAAGAGAGAAAGCAGAGCTGTCTGATCATCTGA 1409
QY 106 -----MetAlaIaAsnSerThrAlaValAlaGluAspIleThrIysAspGly 121
Db 1410 CCTGCAGAGCTTCATCTCCAACTTCCAGCCAGCCAGAGAGAGCTGAGCAGGCTCCCA 1469
QY 121 nGluIuThrSerGluIleIleGlnGlnIleProAlaSerGluIaAsnValGluIuIe 141
Db 1470 GGCTCATGGGGCCGCTTACTGCTGAGTGAGCTCTCTGACCTGTGAGCTCACCACT 1529
QY 141 tValGlnProAlaGluSerGlnAlaAsnAspValGlyPheIysIysValPheIysPhe 161
Db 1530 CAATGCCACCATCCAGCAACAGATCAAGAACTG-----GC 1565
QY 161 lGlyPheIysPheThrValIyIyAspIyAsnGlnIySeryAspThrValGlnIe 181
Db 1566 TGGCTTGAAGCAGCAGGCCAAAGAGAGAGCCAGCTAGCAGACCTTCCAA----- 1620
QY 181 uThrValIyIySeryAspGlnGluIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 201
Db 1621 -----CAGCAAGAAACAGGCTCTCCAGGGCTTC-----CGCCACCA 1655
QY 201 nGluProSerValGluThrAlaValGlyIuSerAlaIySeryGlnIuIeIyG 221
Db 1656 G-----GTGAGAGCAGCTAAGCAGTAGCCTGAAGCAGAGAGCAGCTGAAGAGA 1706
QY 221 nSerThrGluIySglnGluIyThrIeIySglnGlnIuIeIySeryThrGlnIlePro 241
Db 1707 GGTAGCGAGAGAGAGAGGAGCACT-----AGCAGAGACCATGCCAGCACTGGCAC 1760
QY 241 uGlnIaIaGluSerAspGlnAlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 259
Db 1761 TGCTGCAAGGAGAGCAGAGGCTCTTAAGGAGAGCGGATGGCTTCAAGCAGCTGGA 1820
QY 259 uIySglnGluIySglnProIuThrIySeryProIuIy-----SerProSerProVal 277
Db 1821 GGCACCTGAGAGAGAGAGAGGCTGCCAAGCTGAGAAATCTGACAGCAACTTCAGGTGC 1880
QY 277 nSerGluThrThrSerSerPheIySeryPheThrIleIyIaIaIaIaIaIaIaIaIa 297
Db 1881 TAATGAAGCCCGGAGCACTGCCACACTCACTGACACAG-----GCCACCGGGA 1931
QY 297 sIySerThrSerPheIySerySeryIyAsp-----AspIeGlnIuThrAlaGluIy 314
Db 1932 GAAGCAGAGCTGACCGGAGAGAGTGAAGAACTCCAGGCTGTGTGAGACAGCCGCA 1991
QY 314 sArgIySglnGlnIuIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 331
Db 1992 GGAAACAGCATGAGGCCCAAGGCCAGTTGACAGAGCTTGAAGCTGCGGCTGAGCA 2051
QY 331 rGluProAlaSerGluIuIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 351
Db 2052 GCAAAAAGCACTTGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2105
QY 351 aAspTyrgIuIySeryValGluIeProIeGluIaIaIaIaIaIaIaIaIaIaIaIa 371
Db 2106 GCAGCTCCAGGCTTCAAAAGAGTCTTGAAGGTGACCAAGGCGACCTTGAA----- 2157

```

QY 371 rGluGluLysCysAlaProLeuAlaThrGluValPheAspGluLysMetGluAlaHisG1 391
Db 2158 -GAGGAGAAG-----CGCAGGGCTGCAGATGCTCCGAAAGACAGCAGGCGTTGTATCTC 2210
QY 391 nGluValValAlaGluVal-----HisValSerThrValGluLysTh 405
Db 2211 TGAGCTGAAGCGCAGAGACCCGAGAGCTGGTGGAGCAGCATAAAGCGGAACGAAAGGAGCT 2270
QY 405 rGluGluGluGlnGlyGlyGlyGluAlaGluGlyValValValGluGlyThrG1 425
Db 2271 GGAAGAAGAGAGAGGCTGGCGCAAGGGCGCTGAGGCTCGATTACTGCAGCTTGGGAGGC 2330
QY 425 yGluSerLeuProGlu-----LysLeuAlaGluProGlnGluValProG1 441
Db 2331 CCATCAGCTGAGACTGAAGTCTCGCGCGGAGCTGGCAGAGCCATGCTGCCAGCA 2390
QY 441 nGluAlaGlu---ProAlaGluLeuMetLysSerArgGluMetCysValSerGlyG1 460
Db 2391 CACAGCTGAGAGTGAGTGTGAGCAGCTCGTCAAGAAGTAGCTGCTGGCGTGACGGTA 2450
QY 460 yAspHisThrGlnLeuThrAspLeuSerProGluGluLysThrLeuProLysHisProG1 480
Db 2451 TGAGGATAGCCAG-----CAAGAGAGGCC-ACAGTATGGCGCCATGTTC 2494
QY 480 uGlyIleValSerGluValGluMetLeuSerSerGlnGluArgIleLysValGlnGlySe 500
Db 2495 AGGAACAGCTGATGATTTGAA----- 2516
QY 500 rProLeuLysLysLeuPheSerSerSerGlyLeuLysLysLeuSerGlyLysLysGlnLy 520
Db 2517 -----GGAGAATGTGAGA 2530
QY 520 sGlyLysArgGlyGlyGlyGlyAspGluGluProGlyGluTy-rGlnHisIleHisThrG1 540
Db 2531 AGGCCGCCAGGAGCTGCAGAGGCAAGGAGAGGTGGCGCAGCATAGAAATCCACAGCG 2590
QY 540 uSerProGluSerAlaAspGluLysGlyGlyLysSerAlaSerSerProGluGluPr 560
Db 2591 AGTCCCA-----GATAAGCCGCGCAGCAGCAACAATAGTGTAGCTCCATGCCAAC 2641
QY 560 oGluGluThrThrCysLeuGluLysGlyProLeuGluAlaProGlnAspGlyGluAla-- 579
Db 2642 TGCCAGAGCACT-----CCAGCAGTCCAGAGAAGGAGTCAAGGCCCGCCAGAGTTG 2695
QY 580 -----GluGluGly-ThrThrSerAspGlyGlyLysLysA 591
Db 2696 CAGATGACCTCTCCACTCTGCAGGAAAGATGGTGCACCACGCAAGAGAGGTGCCCGCT 2755
QY 591 rGluGlyIleThrProTrpAlaSerPheLysLysMetValThrProLysLysArgValA 611
Db 2756 TGGAGACCTTGGTCCGCAAGGCAAGTGTGAGCAGCAGGAAACAGAGCTCCCGGGAGTTAGTCA 2815
QY 611 rArgProSerGluSer---AspLysGluGlu-----LeuGluLysValLysSerAlaT 629
Db 2816 AGGAGCTCGAGGCGAGGAGACAGCAGCCGAGTGGTGGGAAGAGCAACAGGGAGCGCC 2875
QY 629 hrLeuSerSerThrAspSerThrValSerGluMetGlnAspGluValLysThrValGlyG 649
Db 2876 AGTTCTGCACACACAGCGCAGCTGCAGGCTATGGAGCGGGAGGACAGCAGATGGGCA 2935
QY 649 luGlu-----GlnLysProGluGlu- 655
Db 2936 ATGAGCTGGAAACGCTGCGGGCGCGCTGTATGAGAGCCAGGGGCGCAGCAGGAGGAGC 2995
QY 655 ----- 655
Db 2996 GTGGGCGAGCGAAGGAGGTGGCGGCTGTACCCAGAGCGGGCGCTGCCAGGCTG 3055
QY 656 -----P 656
Db 3056 ACCTTGCCCTGGAGNAGCGCGCAGCAGAGCTTGAGATGCGGCTGCAGAACGCCCTCA 3115
QY 656 roLysArgArgValAspThrValSerValSerTrpGluAlaLeuIleCysValGlySerSerL 676

Db 3116 ACAGCAGCGTGTGGAGTTGCTTACCTGCAAGAGCACTGGCTCATGCCCTCACGGAAA 3175
QY 676 yLysArgAlaArgLysAlaSerSerSerAspAspGluGlyGlyProArgThrLeuGlyG 696
Db 3176 AGGAA-----GGCAAGGACCGAGGTT-GGCCAAGGCTTCGTGGTCTG 3216
QY 696 lyAspSer-----HisArgAlaGluGluAlaSerLysAspLysGluAla-----GlyT 712
Db 3217 GAGCGAGCCAGATAAAGAGCTGGAGGAACCTTCGGCAAAACCGTGAAGCAACTGAAGAA 3276
QY 712 hrAspAlaValProAla-SerThrGlnGluGlnAspGlnAlaGlnGlySerSerSerPro 731
Db 3277 CA-----GCTGGCTAAGAAAGAAAGGACGACCATCTGCTCAGAGGCCAA 3324
QY 732 GluProAlaGlySerProSerGluGlyGluGlyValSerThrTrpGluSerPheLysArg 751
Db 3325 TCTGAGGCTCTGGCAGGACAGAGCCAAACAGGCCCAAGCTG---GAAGCACTGGGGCA 3381
QY 752 LeuValThrProArgLysLysSerLysSerLysLeuGluLysAlaGlu----- 768
Db 3382 GAGCTGAGCAAGCTGGAACAGCAATGCCAAGAGCAGCAGCAGCAGCTGACAGCTGGAA 3441
QY 769 -----AspSerSerValGluGlnLeu 775
Db 3442 CGCAGCTCGAGGCTGAGCGGGCTCCCGGGCTGAGCGGACAGTCTCTGGAGACTCTG 3501
QY 776 SerThrGluIleGluProSerArgGluGluSerTrpValSerIleLysLysPheIlePro 795
Db 3502 CAGGGCCAGTAGAGAGAGAGGCCAGGAGCTAGGCCACAGTCAGAGTGCCTTAGCCTCG 3561
QY 796 GlyArgArgLysLysArgAla-----AspGlyLysGlnGlnAlaThrValGluAsp 813
Db 3562 GCCCAACGGGAGTTGGCTGCTTCCGCACCAAGGTACAAGACCACACAGAGGCTGAAGAT 3621
QY 814 SerGlyProValGluIle-----AsnGluAspAspProAsnValProAlaVal 829
Db 3622 GAGTGAAGAGCCCGAGTGGCGGCCCGCGCAAGAGGCTGAGAGAGAAAATAGCCTCATC 3681
QY 830 ValProLeu---SerGluTyrrhenAlaValGluArgGluLysMetGluAlaGlnGlyAsn 848
Db 3682 AGCAGCTTGAGGAGAGAGGTGTCCATCTGTAATCGCAGGCTCTGGAGAGGAGGGG--- 3738
QY 849 ThrGluLeuProGlnLeuGlyAlaValTyrrValSerGluGluLeuSerLysThrLeu 868
Db 3739 ---GAGACCAAGGAGTTGAAGCGGTGGTGTGATGGCCGAGTCAGAGAGAGCCAGAGCTG 3795
QY 869 ValHisThrValSerValAlaValIleAspGlyThrArgAlaValThrSerValGluGlu 888
Db 3796 GAGGAGAGCTGCGCTGCTGC-----AGGCAG 3822
QY 889 ArgSerProSerTrpIleSerAlaSerValThrGluProLeuGluHisThrAlaGlyGlu 908
Db 3823 AGACACCCAGCA----- 3834
QY 909 AlaMetProValGluGluValThrGluLysAspIleAlaGluGluThrProVal 928
Db 3835 -----ACAGTGCAGAGCTGCAGAGACGAGCTCTGCTCTGCGGGAGG----- 3876
QY 929 LeuThrGlnThrLeuProGluGlyLysAspAlaHisAspAspMetValThrSerGlu--- 947
Db 3877 -----AGGTGCAGAGCTCCGGGAGGAGGCTGAGAAAACAGCGGTGGCTTCAGAGAAC 3930
QY 948 -----ValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeuArg 965
Db 3931 CTGCGCAGAGAGTGTCTTCAGAAAGAGCAGGCGCTCTCCACCTCGAGCTC----- 4044
QY 966 Thr-----GluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSer 983
Db 3991 GCGTGCCAGAGAGAGTTCTTCAGAAAGAGCAGGCGCTCTCCACCTCGAGCTC----- 4044
QY 984 AlaValSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGlu 1003

Db 4045 -----GAGCACACGACACAGGCC 4065
 Qy 1004 ValGluSerGlyValLeuAspThrGluGluGluValArgGlnThrGlnAlaIleuGln 1023
 Db 4066 CTGGTAGAGTGAAGCTGCTGACAGTAAGCACTCTGACAGCACTGACAGCCAGCAGGCC 4125
 Qy 1024 AlaValAlaAspLeuValLeuGluGluSerGlnValProAlaThrGlnThrValGlnArg 1043
 Db 4126 GTGGCCGAGAAAGCCGCTGAGAGCTGGAG-----CAGAGCAAGAGGCC 4173
 Qy 1044 ThrGlySerGlyAlaLeuGlu----- 1050
 Db 4174 GTGGGGAGCATCCGGGAGAGCTGCTGGGGCCAGCGGAGCTTGGGGAGCTGATCTCT 4233
 Qy 1051 -----LysValGluGluValGluGluAspSerGluValLeuAlaSerGlyLys--- 1066
 Db 4234 CTGGCAGCAAGAGTGGCAGAGCAGAGCGAACAGCTCAGAGCTGGCGGAGAGAGGCC 4293
 Qy 1066 ----- 1066
 Db 4294 AGCTATGACAGACGACTGAGCATGCTGAAGAAAGCGCATGGCTGCTGGCAGAGAGAAC 4353
 Qy 1066 ----- 1066
 Db 4354 CGGGGGCTGGTGAGCGGGGCCAACTTGGCCGAGCTTCTGGAAGTGAGATTGACACAG 4413
 Qy 1067 -----GluLysAspValMetProLysGlyProValGlnGluAlaGlyValaGluHisLeu 1084
 Db 4414 GCCCGGAAAGATGATGTCCAAAGATTGGCAGCGCTGACGTGATGCTAGACCCCTGTG 4473
 Qy 1085 AlaGlnGlySerGluThrGlnAlaThrProGluSerLeuGluValProGluValThr 1104
 Db 4474 GGTGAGTGACACGAGACACAGACGACTGCCGGGAGCTGAGAGTG-----ATGACT 4527
 Qy 1105 AlaAspValAspHisValAlaThrCysGlnValIleLysLeuGlnLeuMetGluGln 1124
 Db 4528 GCCAAGTATGAGGCTCC-----AAGGTCAAGCTCTTGAGAGAGAG 4569
 Qy 1125 AlaValAlaProGluSerGluThrLeuThrAspSerGluThrAsnGlySerThrPro 1144
 Db 4570 CAGCGGTCCAGAGAGAGAGAGAACTCAGTGCAGAGTGGAGAAACTGATGAGAA 4629
 Qy 1145 LeuAlaAspSerThrAlaAspGlyThrGlnGlnAsp-----GluThrIleAspSer 1162
 Db 4630 CTGGCTGACTCTGACCAACGACGAGGTGACAGCAGAAAGCTGAGCGCTGACAGGT 4689
 Qy 1163 Gln-----AspSerValAlaThrAlaAlaValArgGlnSerGlnValThrGluGluGln 1180
 Db 4690 CAGGAGGCGAGAGCCAGAGAGAGGCCCTTCCAGGCCCAAGTGAATGAATGCA 4749
 Qy 1181 AlaAlaThrAlaGlnLysGluGluProSer-----ThrLeu 1192
 Db 4750 GCCCAGTTGAGCCAGAGAGAGAGCGAGCACTATAGCTGAGATGAGAAAGCC 4809
 Qy 1193 ProAsnAsnValProAlaGlnGluGluHisGlyGluGlu-----ProGly 1207
 Db 4810 AAAACACATTATGATGCCAAGAGCAGCAGAACCAAGACTGACAGAGAGAGCTGGAGC 4869
 Qy 1208 ArgAspValIleuGluProThrGlnGlnGluLeuThrAlaAlaValProValIleuAla 1227
 Db 4870 CTGGAGCACTGACAGAGAAACAAAGAGCTGCGAGCTGAACT----- 4914
 Qy 1228 LysThrGluValGlyGlnGluGluValAspTrpLeuAspGlyGluLysValLysGlu 1247
 Db 4915 ---GAACGCTGGGCGCATGACTACAG-----CAGGCTGGCTGAAGACCAAGAG 4962
 Qy 1248 GluGlnGluValaPheValHis-----SerGlyProAsnSerGlnValaAlaAspVal 1265
 Db 4963 GCTGAACACACTGCGCGCACTTACTGCTCAGGTCGAGCAGCTGGAGCAGAGTTGCC 5022
 Qy 1266 ThrTyrAspSerGluValMetGlyValaGlyCysGln-----GluLysGlu 1281
 Db 5023 CATGACAGACCAAGCTTGAAGACTGGGCAATTCCAGGTGGCAACTGATGCTTTAAAG 5082

Qy 1282 SerThrGluValGlnSer-----LeuSer 1289
 Db 5083 AGCCCTGAGCCCCAGAGCTAAGCCCCAGCTGAGCTTGAAGTACAGCCCTGATCTGAGC 5142
 Qy 1290 LeuGluGluGlyLysMetGluThrAspValGluLysGlyLysArgGluThrLysProGlu 1309
 Db 5143 TGGAGAGGGAGAGCCCACTAGTATACAGCAAGAGTGCCTGT--ACCAGCCAGAAC 5199
 Qy 1310 GlnValSerGluGluGlyGluGlu-----ThrAlaAlaProGluHisGluGly 1326
 Db 5200 GGCACCAAGCTCCCTGAGAGAACAGAGCTCAGCTATCTCCA-GCGCTGCCCCCAAGT 5258
 Qy 1327 ThrTyrGlyLysProValLeuThrLeuAspMetProSerSerGluValaValaLeu 1346
 Db 5259 AGAATCCCTGAGAGAGTCTTACTTCACTCCATCCAGCTGCTGGAGTCCAGGCCCTCT 5315
 Qy 1347 GlySerLeuGlyGlySerProSerLeuPro----- 1356
 Db 5316 -----GAGAGCAGCTGAGCTCCCTGGAGAGCTTCTTGACCTGGCTGTAA 5366
 Qy 1357 -----AspGlnAspLys 1360
 Db 5367 GACCCGCTCCGCTGCGGCGACACGACAGATCATCAACATCAACATCAACAGAA-- 5423
 Qy 1361 AlaGlyCysIleGluValGlnValGlnSerLeu-AspThrValThrGlnThrAlaGln 1380
 Db 5424 GCTAGATGTGAAGAGCCAGACAGCGCACTCATGTTTACAGCAGCGGTGTCTCC 5483
 Qy 1380 AlaValGluLysValIle-----GluThrValValIleSerGluThrGlyGln 1396
 Db 5484 TGCTTCCAGGCTAGCTGACGAGCCAGCTCTCTACGATCTAGCTCCGTGGGT-- 5541
 Qy 1396 userProGluCysValaGlyAlaHisLeuLeu----- 1406
 Db 5542 -TCTCCGATTATGCACTCAGCCCTGCTCAGCTTGCCTGAGTACGCCCAACCACTCG 5600
 Qy 1407 ---ProAlaGluLysSerSerAlaThrGlyGlyHisIleThrLeuGlnHisAlaGluAs 1425
 Db 5601 CAGTCTCTGCTGCTGCTCCAGGCGCG--GATGCCAGTGGGCCCCCTCAGAGAAAGACA 5659
 Qy 1425 PThrValPro-----LeuGlyProGluLeu 1433
 Db 5660 GCTTTTACATGGGCACTTGCAGAGTAGAGCTGAGCAGCTGAGTACGTGAACCCGATTG 5719
 Qy 1433 rGlnAlaGluSerIleProIleIleValThrProAlaProGluSerThrLeuHisProAs 1453
 Db 5720 CAGAGCTGACAGAGCGCAATGAGTGTCCCCCAGATGAAAGCCTATGCCCTGG 5779
 Qy 1453 pLeuGlnGlyGluIleSerAlaSerGlnArgGluArg--SerGluGluGluAspLysPr 1472
 Db 5780 AGTCCAGGCTTCCCTGAGCCTGGCAGCAACATCAAGAGAGAGTGAAGAACTGAGAAC 5839
 Qy 1472 oAsp-----AlaGlyProAspAlaAspGly 1480
 Db 5840 CCAAGAGACCTGCGCCGAGCAGACATGACCAATCCAGATAGCCGAGGG 5891

RESULT 11
 US-08-195-487-3
 ; Sequence 3, Application us/08195487
 ; Patent No. 5783403
 ; GENERAL INFORMATION:
 ; APPLICANT: TONKATLY, GARY
 ; APPLICANT: LIDGARD, GRAHAM P
 ; TITLE OF INVENTION: NOVEL MALIGNANT CELL TYPE MARKERS OF THE
 ; TITLE OF INVENTION: INTERIOR NUCLEAR MATRIX
 ; NUMBER OF SEQUENCES: 6
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: TESTA HURWITZ & THIBEAULT
 ; STREET: 53 STATE STREET
 ; CITY: BOSTON
 ; STATE: MA
 ; COUNTRY: USA

ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/195,487
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/901,701
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: PITCHER ESQ, EDMUND R
REGISTRATION NUMBER: 27,829
REFERENCE/DOCKET NUMBER: MTP-013
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/248-7000
TELEFAX: 617/248-7100
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 6306 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..6306
PUBLICATION INFORMATION:
AUTHORS: COMPTON, DUANE A
AUTHORS: SZILAK, ILLYA
AUTHORS: CLEVELAND, DON W
TITLE: PRIMARY STRUCTURE OF OF NUMA, AN INTRANUCLEAR
TITLE: PROTEIN THAT DEFINES A NOVEL PATHWAY FOR
TITLE: SEGREGATION OF PROTEINS AT MITOSIS
JOURNAL: JOURNAL OF CELL BIOLOGY
VOLUME: 116
PAGES: 1395-1408
DATE: MAR-1992
US-08-195-487-3

Alignment Scores:
Pred. No.: 3,2e-19 Length: 6306
Score: 412.00 Matches: 372
Percent Similarity: 35.28% Conservative: 256
Best Local Similarity: 20.90% Mismatches: 663
Query Match: 5.10% Indels: 493
DB: 1 Gaps: 72

US-09-902-432-4 (1-1596) x US-08-195-487-3 (1-6306)

QY 28 GlyHisGlyProAlaAlaGluAlaSerGlyAlaAlaGlyAspProAlaAspAlaAspPro 47
Db 1050 GGCACCTCAGAGTGGCTAGAGAGAGAGGCGCCAGCTGGAGAGGAGCTCAGCGCAGCCCT 1109
QY 48 Ala-ThrLys-----LeuProGlnLysAsn-----GlyGlnLeuSerSerVa 61
Db 1110 GCAGGACAGAAATGCTTGNAGAGAGAGACGAATCCTTCAGGGAANAATCTTCACAGCT 1169
QY 61 l-----AsnGlyValAlaAlaGlnGlnGlyAspVal----- 70
Db 1170 GGAAGAACACTTGTCCAGCTGCAGGATAACCCACCCAGGAGAGGCGGAGGTGCTGGG 1229
QY 71 -----HisValGlnGlnGluAsnGlnGlnGly----- 79
Db 1230 TGATGTCTTGAGCTGGAAACCTTTGAAGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1289
QY 80 -----GlnGlnGlnGlnValValAspGluAspValGlyGlnArgGluSerGluAs 96
Db 1290 ACAGCTCAAGCCAGGGTAGAGATGCTGGAGACTGAGCGAGGCGCAGGAGGAGGAGGAGGAG 1349

QY 96 pValArgGluLysAspArgValGluGlu----- 105
Db 1350 GCTTGTGAGCGGGGCCACTTTCGAAGAAGAAAGACAGCAGCTGTCTAGCTGTGATCACTGA 1409
QY 106 -----MetAlaAlaAsnSerThrAlaValGluAspIleThrLysAspGlyGly 121
Db 1410 CTTGCAGAGCTCCATCTCCAACTCAGCCAGGCGCAAGAGAGAGCTGGAGCAGCCTCCCA 1469
QY 121 nGluGluThrSerGluIleGluGlnIleProAlaSerGluAsnValGluGluMe 141
Db 1470 GGCTCATGGGGCCGGTGTGACTGCCAGGTGGCTCTCTGACCTCTGAGCTCACCACACT 1529
QY 141 tValGlnProAlaGluSerGlnAlaAsnAspValGlyPheLysLysValPheLysPheVa 161
Db 1530 CAATGCCACCATCCAGCAACAGGATCAAGAACTG-----GC 1565
QY 161 lGlyPheLysPheThrValLysLysAspLysAsnGluLysSerAspThrValGlnLeuLe 181
Db 1566 TGGCCTGAGCAGCAGCGCCAAAGAGAGCAGCGCCAGCTAGCACAGACCTCCAA----- 1620
QY 181 uThrValLysLysAspGluGlyAlaGluAlaSerValGlyAlaGlyAspHisGly 201
Db 1621 -----CAGCAAGAACAAGCGCTCCAGGGCCTC-----CGCCACCA 1655
QY 201 nGluProSerValGluThrAlaValGlyGluSerAlaSerLysGluSerGluLeuLysGly 221
Db 1656 G-----GTGGAGCAGCTAAGCAGTAGCTGAAAGCAGAGGAGCAGCAGGTGAAGGA 1706
QY 221 nSerThrGluLysGlnGlyThrLeuLysGlnGlnSerSerThrGluIleProLe 241
Db 1707 GTTAGCGGAGAGCAGGAGCAACT-----AGCAGGACCATTGCCAGCAACTTGGCCAC 1760
QY 241 uGlnAlaGluSerAspGlnAlaAla-----GluGluGluAlaLysAspGluGlyGluGly 259
Db 1761 TGCTGCAGAGGAGCGAGAGGCTCTTAAGGAGCGGATGCGGCTCTCAAGCAGCTGGA 1820
QY 259 uLysGlnGluGluProThrLysSerProGlu-----SerProSerSerProValas 277
Db 1821 GGCACTGGAGAGAGAGAGCTGCCAAGCTGGAGATTCTGCAGCAGCAACTTTCAGGTGGC 1880
QY 277 nSerGluThrThrSerSerPheLysLysPhePheThrHisGlyTrpAlaGlyTrpArgly 297
Db 1881 TAATGAAGCCCGGACAGTCCCGCAGACCTCAGTCAGCAG-----GCCAGCGCGGA 1931
QY 297 sLysThrSerPheLysLysSerLysGluAsp-----AspLeuGluThrAlaGluLys 314
Db 1932 GAAGGCAGAGCTCAGCGGAGAGTGGAGGAACCTCCAGGCTGTGTGAGACAGCCGCCCA 1991
QY 314 sArgLysGluGlnGluAlaGluLysValAspGluGluGlu-----LysGluLysTh 331
Db 1992 GGAACAGCATGAGGCCCGCAGGCTTGCAGAGTAGAGTTGCAGCTGCGGTCTGAGCA 2051
QY 331 rGluProAlaSerGluGluGlnGluProAlaGluAspThrAspGlnAlaArgLeuSerAl 351
Db 2052 GCAAAAGCAACTGAGAAAGAGGTTGGCCCGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 2105
QY 351 aAspTyrGluLysValGluLeuProLysGluAspGlnValGlyAspLeuGluAlaSerSe 371
Db 2106 GCAGCTCCAGGCCCTCAAGAGTCTTGAAGGTCCCAAGGCGCAGCTTGAA----- 2157
QY 371 rGluGluLysCysAlaProLeuAlaThrGluValPheAspGluLysMetGluAlaHisGly 391
Db 2158 -GAGGAGNAG-----CGCAGGGCTGCAGATGCTTGAAGAGCAGCAGCGGCTGTATCTC 2210
QY 391 nGluValValAlaGluVal-----HisValSerThrValGluLysTh 405
Db 2211 TGAGCTGAAGGCGAGAGACCCGAAAGCTGTGGAGCAGCATTAAGCGGGAACAAAGAGGACT 2270
QY 405 rGluGluGluGlnGlyGlyGlyGlyGluAlaGluGlyGlyValValValGluGlyThrGly 425
Db 2271 GGAAGAGAGAGGCTGGCGCGAAGGGGCTGGAGGCTCGATTACTCGAGCTTGGAGGCG 2330
QY 425 yGluSerLeuProProGlu-----LysLeuAlaGluProGlnGluValProGly 441

```

Db 2331 CCATCAGCTGAGACTGAATCTCTGCGGGGAGCTGGCAGAGGCCATGCTGCCAGCA 2390
Qy 441 nglualaglu---ProLaIgluIleuMetLysSerArgIleuMetCysValSerGly 460
Db 2391 CACAGCTAGAGTGTGTGTGAGACACTCGTCAAGAAGTAGCTGCTGGCGTGAAGGGGA 2450
Qy 460 yAspHisThrGluLeuThrAspLysSerProGluGluIlyThrLeuProLysHisProG 480
Db 2451 TGAGGATGAGCCAG-----CAAGAGAGAGG-ACAGTATGGCGCGCAATGTTCC 2494
Qy 480 uGlyIleValSerGluValGluMetLeuSerSerGlnIleuArgIleuValGlnGlySe 500
Db 2495 AGGAACAGACTGATGCTTTGAA----- 2516
Qy 500 rProleuLysLysLeuPheSerSerSerGlyLeuLysLysLeuSerGlyLysGlnLys 520
Db 2517 -----GGAGGAATGTGAGA 2530
Qy 520 sGlyLysArgGlyGlyGlyLysPheGluIleuProGlyIleuHisIleHisThrG 540
Db 2531 AGGCCCGCAGAGAGCTGCAAGAGAGCAAGAGTAGCGCAGCATAGATCCACAGCG 2590
Qy 540 uSerProGluSerAlaAspGluGlnLysGlyLysSerAlaSerSerProGluIleu 560
Db 2591 AGCTCCA-----GATTAAGCCGCGCAGACACAACTAGACTGATGCCATCC 2641
Qy 560 cGluIleuThrThrCysLeuGluIleuGlyProleuGluIleuProGlnAspGlyIleu 579
Db 2642 TGGCCAGAGCACT-----CCAGCAGGTCCAGAGAGAACTCAGGCCCGCAAGGCTTG 2695
Qy 580 -----GluGluGly-ThrThrsSerAspGlyIleuLysAs 591
Db 2696 CAGATGACCTCTCCACTCTGCAGAAAGATGCTGCCACAGCAAGAGTGCGCCGCT 2755
Qy 591 rGluGluGlyIleuThrProTrpAlaSerPheLysLysMetAlaThrProLysLysArgVal 611
Db 2756 TGGAGACCTTGCTGCGCAAGAGAGTGAAGCAGAGAAACAGCTCCCGGAGTATGTA 2815
Qy 611 rGArgProSerGluSer---AspLysGluGluGlu---LeuGluLysValLysSerAla 629
Db 2816 AGGACCTTCGCGAGGAGAGACAGACAGCCGAGTGCCTGGAAGCAAGCAAGGAGCGCC 2875
Qy 629 hrLeuSerSerThrAspSerThrValSerGluMetGlnAspGluValLysThrValGly 649
Db 2876 AGTCTGAGACACAGAGCAGCGCTGAGGCTATGAGCGGAGGAGGAGCAGATGGGCA 2935
Qy 649 IuGluIu-----GluLysProGluGlu- 655
Db 2936 ATGAGCTGGAAAGCTGCGCGCGCGCTGATGAGAAGCCAGGGCGACAGAGAGAGC 2995
Qy 655 ----- 655
Db 2996 GTGGGACAGCAAGAGAGGTGGCGGCTGACCAGAGAGCGGGCCCTGCCAGGCTG 3055
Qy 656 -----P 656
Db 3056 ACCTTGCCCTGGAGAAAGCGCCAGACAGAGCTTGAGATGGCGCTGCAGAAAGCCCTCA 3115
Qy 656 rGlyArgArgValAspThrSerValSerTrpGluAlaLeuIleCysValGlySerSerL 676
Db 3116 ACGACAGAGCTGTGGAGTTGCTACCTGCAAGAGGACATGGCTCATGCCCTGAGGAAA 3175
Qy 676 yLysArgAlaArgLysLysSerSerSerAspAspGluGlyGlyProArgThrLeuGly 696
Db 3176 AGGAA-----GGCAGAGCCAGAGATT-GGCCAAGCTTGCTGCTCG 3216
Qy 696 IYAspSer-----HisArgAlaGluGluAlaSerLysAspLysGluAla-----GlyT 712
Db 3217 GAGGAGGCCCAATTAAGAGCTGGAGGAACTTGGCAAAACCTGTAAGCAACTGAAGGAA 3276
Qy 712 hrAspAlaValProAla-SerThrGlnGluGlnAspGlnAlaGlnGlySerSerSerPro 731

```

```

Db 3277 CA-----GCTGGCTAAGAAAGAAAGAGCAGCATGCTGGCTCAGAGCCCA 3324
Qy 732 GluProAlaGlySerProSerGluGlyGluGlyValSerThrTrpGluSerPheLysArg 751
Db 3325 TCTGAGGCTCTGGCAGAGGACAGAGCCAAAGAGCCCAAGCTG---GAAGCACTCGGCA 3381
Qy 752 LeuValThrProAlaGlyLysLysSerLysSerLysLeuGluGlyLysAlaGlu----- 768
Db 3382 GAGGTAGCAAACTGGAAACAAATGCCAGAAAGCAGAGGAGCTGAGACGCTTGAAA 3441
Qy 769 -----AspSerSerValGluGluLeu 775
Db 3442 CGCAGCTCGAGGCTGAGCGGCGCTCCGCGGCTGAGCGGCACTGCTCTGGAAGACTCG 3501
Qy 776 SerThrGluIleGluProSerArgGluIleuSerTrpValSerIleLysLysPheIlePro 795
Db 3502 CAGGCGCAGTTAGAGAGAAAGCCAGAGCTGAGGCACTGACAGAGGCTTAGGCTCG 3561
Qy 796 GlyArgArgLysLysArgAla-----AspGlyLysGlnGlnGlnAlaThrValGluAsp 813
Db 3562 GCCCAACGGAGATTGCTGCTGCTCCGACCAAGGTACAAAGCACAAGCAGCTGAAGAT 3621
Qy 814 SerGlyProValGluIle-----AsnGluAspAspProAsnValProAlaVal 829
Db 3622 GAGTGGAAAGGCCAGTGGCCCGGCGCGGCGCAAGAGGCTGAGAGAAAAATAGCTCATC 3681
Qy 830 ValProLeu---SerGlyTrpAsnAlaValGluArgGlyLysMetGluAlaGlnGlyAsn 848
Db 3682 AGCAGCTTGAGAGAGAGGTGCTCATCTGTAATCGCCAGGTCCTGGAGAAAGAGGCG--- 3738
Qy 849 ThrGluLeuProGlnLeuLeuGlyAlaValIleLysSerGluGluLeuSerLysThrLeu 868
Db 3739 ---GAGAGCAAGAGTTGAAACGCGCTGTGTATGGCCGAGTGCAGAGAGAGCAGAGCTG 3795
Qy 869 ValHisThrValSerValAlaValIleAspGlyThrArgAlaValThrSerValGluGlu 888
Db 3796 GAGAGAGACTGCGCTGCTGTC-----AGGCAG 3822
Qy 889 ArgSerProSerTrpIleSerAlaSerValThrGluProleuGluHisThrAlaGlyLys 908
Db 3823 AGACAGCCAGCA----- 3834
Qy 909 AlaMetProProValGluGluValThrGluLysAspIleIleAlaGluGluThrProVal 928
Db 3835 -----ACAGTGCAGAGCTCGAGAACCGACCTGCTCTCGGAGG----- 3876
Qy 929 LeuThrGlnThrLeuProGluGlyLysAspAlaHisAspAspMetValThrSerGlu--- 947
Db 3877 -----AGGTGCAAGACTCCGGAGGAGAGGCTGAGAAACAGCGGCTGCTTCAGAGAAC 3930
Qy 948 -----ValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeuArg 965
Db 3931 CTGGCGCAGAGAGCTGACTCAAGGCTGAGCGTGGAGAGAGCTGGCGCCAGAAATTGAAG 3990
Qy 966 Thr-----GluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSer 983
Db 3991 GCGTGGCAGAGAGATTCTTCAGAAAGAGCAGGCTCTCCACCTCCAGCTC----- 4044
Qy 984 AlaValSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGlu 1003
Db 4045 -----GAGCAGCAGCAGCAGCAGGCC 4065
Qy 1004 ValGluSerGlyValLeuAspThrGluGluGluGluGluGluArgGlnThrGlnAlaIleLeuGln 1023
Db 4066 CTGCTGATGAGTACTGCTCCAGCTAAGCACTCTGCCAGAGCTGCAGAGCCGAGAGAGGCC 4125
Qy 1024 AlaValAlaAspLysValLysGluGluSerGlnValProAlaThrGlnThrValGlnArg 1043
Db 4126 GCTGGCGAGAAACGCCACCGAGGAGAGCTGGAG-----CAGAGCAGAGAGGCC 4173
Qy 1044 ThrGlySerLysAlaLeuGlu----- 1050
Db 4174 GCTGGGGAGACTGGCGGAGAGACTGCTGCGGGCCAGCGGAGAGCTTGGGAGACTGATTCT 4233

```



```
QY 1051 -----LysValGluGluValGluGluAspSerGluValLeuAlaSerGluLys--- 1066
Db 4234 CTGCGGAGAAAGTGGCAGAGCAGGACCGAACAGCTCAGCAGCTGCGGAGAGAGGCC 4293
QY 1066 ----- 1066
Db 4294 AGCTATCAGAGCAGCTGAGCATGCTGAAGAGCGCATGGCTGTGGCAGAGGAGAAC 4353
QY 1066 ----- 1066
Db 4354 CGGGGGCTGGTGGCGGGCCAACTTGGCGGCGAGTTTCTGGAAGTGGAGTTGGACCAG 4413
QY 1067 -----GluLysAspValMetProLysGlyProValGlnGluAlaGlyAlaGluHisLeu 1084
Db 4414 GCCCGGAAAAGTATGTCGAAGAGTTGGCAGCGCTACGTCTGATGCTGAGACCCGCTG 4473
QY 1085 AlaGlnGlySerGluThrGlyGlnAlaThrProGluSerLeuGluValProGluValThr 1104
Db 4474 GCTGAGGTGAGCGAGNAGCAGAGAGCAGTCCCGGAGCTGGAGTG-----ATGACT 4527
QY 1105 AlaAspValAspHisValAlaThrCysGlnValIleLysLeuGlnLeuMetGluGln 1124
Db 4528 GCCAAGTATCAGGTGCC-----AAGGTCAAGGTCTCTGGAGGAGG 4569
QY 1125 AlaValAlaProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrPro 1144
Db 4570 CACCGGTTCCAGGAAGAGAGCGCAAACTCACTGCCAGGTGGAAGAACTGAGTAAGAAA 4629
QY 1145 LeuAlaAspSerAspThrAlaAspGlyThrGlnGlnAsp-----GluThrIleAspSer 1162
Db 4630 CTGGTGACTTCTACCAAGCAGCAGAGGTGCGCAGAGCAGAGCTGAAGGCTGTCCAGGCT 4689
QY 1163 Gln-----AspSerLysAlaThrAlaAlaValArgGlnSerGlnValThrGluGluGlu 1180
Db 4690 CAGGGAGGCGAGCCAGCAGGAGGCCAGCGCTTCAGGCCCCAGCTGAATGAAGTCAA 4749
QY 1181 AlaAlaThrAlaGlnLysGluGluProSer-----ThrLeu 1192
Db 4750 GCCCAGTTGAGCCAGAGGAGCAGCGCAGCTGAGCACTATAAGCTGCAGATGGAGAAAGCC 4809
QY 1193 ProAsnValProAlaGlnGluGluHisGlyGluGlu-----ProGly 1207
Db 4810 AAACACATTATGATGCCAAGAGAGCAGCAGAACCAAGAGCTGCAGGAGCAGTCCGAGC 4869
QY 1208 ArgAspValLeuGluProThrGlnGlnGluLeuThrAlaAlaValProValLeuAla 1227
Db 4870 CTGGAGCAGCTGCAGAGGAAACAAAGAGCTGCGAGCTGAAGCT----- 4914
QY 1228 LysThrGluValGlyGlnGluGlyGluValAspTrpLeuAspGlyGluLysValLysGlu 1247
Db 4915 --GAACGGTGGGCCATGAGCTACAG-----CAGGCTGGCTGGAAGACCAAGAG 4962
QY 1248 GluGlnGluValPheValHis-----SerGlyProAsnSerGlnLysAlaAlaAspVal 1265
Db 4963 GCTGAACAGACCTGCCGCCCACTTACTGCCAGGTGGCAGCTGGAGGCAGAGTTGCC 5022
QY 1266 ThrTyAspSerGluValMetGlyValAlaGlyCysGln-----GluLysGlu 1281
Db 5023 CATGCAGACCAGCAGCTTCCAGACCTGGGCAAAATTCAGGTGGCAACTGATGCTTTAAG 5082
QY 1282 SerThrGluValGlnSer-----LeuSer 1289
Db 5083 AGCGGTGAGCCCGGCTAAGCCAGCTGGACTTGAGTATTGACAGCCTGGATCTGAGC 5142
QY 1290 LeuGluGluGlyGluMetGluThrAspValGluLysGluLysArgGluThrLysProGlu 1309
Db 5143 TGCAGAGAGGGAGCCCACTCACTATATCACCAGCAAGCTGCTGT---ACCCAGCCAGAC 5199
QY 1310 GlnValSerGluGlyGluGlnGlu-----ThrAlaAlaProGluHisGluGly 1326
Db 5200 GGCACACAGCTCCTGGAGAACCAAGCCTCACTATCTCCCA-CGGCTTGGCCCCCAGGT 5258
```

```
QY 1327 ThrTyGlyLysProValLeuThrLeuAspMetProSerSerGluArgGlyLysAlaLeu 1346
Db 5259 AGAATCCCTGGAGAGTCTCTACTTCACTCCCATCCCTGCTCGGAGTCAGGCCCCCT-- 5315
QY 1347 GlySerLeuGlyLysSerProSerLeuPro----- 1356
Db 5316 -----GGAGAGCAGCTGGACTCCCTGGGAGAGCGTCTTCTGAGCTCGGGTCGTAA 5366
QY 1357 -----AspGlnAspLys 1360
Db 5367 GACCCGCTCGCTCGTGGCGCACCGAGATCATCAATCACCATCAGCAAGAA--- 5423
QY 1361 AlaGlyCysIleGluValGlnValGlnSerLeu-AspThrThrValThrGlnThrAlaGln 1380
Db 5424 GCTAGATGTGGAAGACCAAGACAGCAGCGCAACTCATCTGTTTACAGCAGCGGGTCTGCTCC 5483
QY 1380 uAlaValGluLysValIle-----GluThrValIleSerGluThrGlyGln 1396
Db 5484 TGGTTCAGGCTAGCTGGAGCCACCTCTCTACTAGTCTCTAGCTCGCCTGGGT-- 5541
QY 1396 userProGluCysValGlyAlaHisLeuLeu----- 1406
Db 5542 -TCTCCGATTATGGCAACTCAGCCCTGCTCAGCTTGGCTACCGCCCCCACCACCTCG 5600
QY 1407 ----ProAlaGluLysSerSerAlaThrGlyGlyHisTrpThrLeuGlnHisAlaGluAs 1425
Db 5601 CAGTTCTGCTCGTCTTCCAGCGCGG-GGTGTCCAGTGGGGCCCTCCAGGAAGGAACA 5659
QY 1425 pThrValPro-----LeuGlyProGluSe 1433
Db 5660 GCTTCTACATGGGCATTGCGAGGATGAGCTGAGCAGCTGGATGACTGGAACCGCATTC 5719
QY 1433 rGlnAlaGluSerIleProIleValThrProAlaProGluSerThrLeuHisProAs 1453
Db 5720 CAGAGCTGCAGCAGCGCAATCGAGTGTGCCCCACCATCTGAAGACCTGCTATCCCTGG 5779
QY 1453 pLeuGlnGlyGluLeuSerAlaSerGlnArgGluArg---SerGluGluGluAspLysPr 1472
Db 5780 AGTCCAGGCTTCCCTGAGCCTGGGCACCATCACAGATGAGGAGATGAAGTGGAGACC 5839
QY 1472 oAsp-----AlaGlyProAspAlaAspGly 1480
Db 5840 CCAAGAGACCTTCCGCCGAGCAGCATGCGACCAATCCAGATAGCCGAGGG 5891

RESULT 12
PCT-US93-06160-3
; Sequence 3, Application PC/TUS9306160
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: NOVEL MALIGNANT CELL TYPE MARKERS OF THE
; TITLE OF INVENTION: INTERIOR NUCLEAR MATRIX
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSER: TESTA HURWITZ & THIBEAULT
; STREET: 53 STATE STREET
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/06160
; FILING DATE: 19930621
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: PITCHER ESQ. EDMUND R
; REGISTRATION NUMBER: 27,829
; REFERENCE/DOCKET NUMBER: MTP-013
; TELECOMMUNICATION INFORMATION:
```

	TELEPHONE: 617/248-7000	
	TELEFAX: 617/248-7100	
	INFORMATION FOR SEQ ID NO: 3:	
	SEQUENCE CHARACTERISTICS:	
	LENGTH: 6306 base pairs	
	TYPE: nucleic acid	
	STRANDEDNESS: single	
	TOPOLOGY: linear	
	MOLECULE TYPE: cDNA	
	FEATURE:	
	NAME/KEY: CDS	
	LOCATION: 1..6306	
	PUBLICATION INFORMATION:	
	AUTHORS: COMPTON, DUANE A	
	AUTHORS: SZILAK, ILITA	
	AUTHORS: CLEVELAND, DON W	
	TITLE: PRIMARY STRUCTURE OF OF NUMA, AN INTRANUCLEAR	
	TITLE: PROTEIN THAT DEFINES A NOVEL PATHWAY FOR	
	TITLE: SEGREGATION OF PROTEINS AT MITOSIS	
	JOURNAL: JOURNAL OF CELL BIOLOGY	
	VOLUME: 116	
	PAGES: 1395-1408	
	DATE: MAR-1992	
	PCT-US93-06160-3	
	Alignment Scores:	
	Pred. No.: 3.2e-19	Length: 6306
	Score: 412.00	Matches: 372
	Percent Similarity: 35.28%	Conservative: 256
	Best local Similarity: 20.90%	Mismatches: 663
	Query Match: 5.10%	Indels: 493
	DB: 5	Gaps: 72
US-09-902-432-4 (1-1596) x PCT-US93-06160-3 (1-6306)		
Oy 28 GLYHISGLYProAlaAlaGluAlaSerGIYAlaAlaGIYAspProAlaaspPro 47		
Dd 1050 GGCACCTCAGAGTGGCTTAAGAAGACAGGCCCACTGGAGAAAGACTTCAGCGCACCCCT 1109	: : : : :	: : : : :
Oy 48 Ala-ThrLys-----LeuProGlnLysasn-----GlyGlnLeuSerSerVa 61		
Dd 1110 GCAGAGACAAGAAATGCTTGAGAGAAAGAACGAATCCTTCAGGGAAAATTCTACACACT 1169	: : : : :	: : : : :
Oy 61 L-----AsnGIYValAlaGluGlnGIYAspVal----- 70		
Dd 1170 GAAGAACACTTGTCCAGCTGCAGATAACCACCCCAGAGAAAGGGCAGGCTCTGGG 1229	: : : : :	: : : : :
Oy 71 -----HisValGlnGlnGluAsnGlnGly----- 79		
Dd 1230 TGATGCTTGCAGCTGGAAACCTTGAAGCAAGAGGACGCCACTCTTGCTGCAACAACAC 1289	: : : : :	: : : : :
Oy 80 -----GlnGlnGluGlnValValaAspGluAspValGIYGlnArgLusSerGluas 96		
Dd 1290 ACAGCTCCAAACCCAGGGTAGAGATGCTGGAGACTGAGCAGCGCAGCAGAAAGCCAAGCT 1349	: : : : :	: : : : :
Oy 96 pValArgGluLysAspArgValGIYglu----- 105		
Dd 1350 GCTTGCTGAGCGGGCCCATCTTCGAAGAAAGAAAGACAGCACTGTCTTACGCTGATCATCTGA 1409	: : : : :	: : : : :
Oy 106 -----MetAlaIlaAsenSerThrAlaValaGluAspLleThrLysAspGlyGI 121		
Dd 1410 CCTGCAGAGCTCCATTCCTCAACCTCACGCCAGGCCAAGAAAGAGCTGGACACAGGCTCCCA 1469	: : : : :	: : : : :
Oy 121 nglunlunhserserluleileleginlinleproalaserGluAsnAsnValGIYglume 141		
Dd 1470 GGCTCATGGGGCCCCGGTTGACTGCCCAAGSGGCGCTCTTGACCTCTGAGACTACACACACT 1529	: : : : :	: : : : :
Oy 141 tvalGlnProAlaGluSerGlnAlaAsnAspValGIYPhelyLysValPhelyPheVa 161		
Dd 1530 CAATGCCACCATTCACGACACAGCAAGGATCAAAGACT-----GC 1565	: : : : :	: : : : :
Oy 161 lGIYPhelyPheThrValLysLysAspLysAsnGlnLysSerAspThValGlnleule 181		

Db	1566	TGGCGCTGAACACACAGCCGCAAAAGAGAGAGAGCCGACCGATCAACAGACCTCCAAA-----	1620
Oy	181	uThrValIyLysAspGlnGlyGlnGlyAlaGlnIAserValGlyAlaGlyAspHisGln	201
Db	1621	-----CAGAGAGAAAGAGGCTCTCCAGAGGCTTC--CGCCACCA	1655
Oy	201	nGluProSerValGluThrAlaValGlyGluSerLaserIyGlnSerGluLeuIySGl	221
Db	1656	G-----GTGGACAGCTAAGACAGTGAAGCTGCTGAAAGCAAGAAAGAGACAGCTTGAAGA	1706
Oy	221	nSerThrGlnIyGlnGlnGlyIThrLeuIySglnGlnGlnSerThrGluIleProLe	241
Db	1707	GGTAGCGGAGAAACAGAGAGCAACT-----AGCAGACACCAATGCCCGCAACTGGCCAC	1760
Oy	241	uGlnAlaGluSerAspGlnAlaAla-----GlnGlnGlnAlaIyAspGlnGlyGlnI	259
Db	1761	TGCTGCAGAGAGACCGAGAGGCTCTTTAAGAGAGGAGATGGAGCTCTCAAGACGTGA	1820
Oy	259	uLyGlnGlnIyLyGlnIyProThrIySerProGlu-----SerProSerProValAs	277
Db	1821	GGCACTGGAAAGAGAAAGAGGCTGCCAAGTGAAGATCTGCAGCAGCAACTTCAAGTGC	1880
Oy	277	nSerGluThrThrSerSerPheIySerIySglnPheThrHisGlyTrpAlaGlyTrpArgIy	297
Db	1881	TAATGAAGCCCCGGAGACAGTCCACAGCTCAAGTGAACAG-----GCCACGGGA	1931
Oy	297	sLyThrSerPheIySerIySerIySglnAsp-----AspLeuGluThrAlaGlnIy	314
Db	1932	GAAAGCAGAGCTGAGCCGAGAGGTGGAGAACTCCAGGCTGTGTGTGAAGACGCCGCA	1991
Oy	314	sArgLyGlnGlnGlnAlaGlnIyValAspGlnGlnGlu-----LysGlnIyIySth	331
Db	1992	GGAACACAGATGAGGCCAGGCCACAGTTGCAGACCTAAGTTGACAGCTGCCTGAGCA	2051
Oy	331	rGluProAlaSerGlnGlnGlnGlnProAlaGlnIySthAspGlnAlaArgLeuSerAl	351
Db	2052	GCAAAAAGCAACTGAGAAAGAAAGAGGTGGCCAGAGAAAGAGCCAG-----CTCCAGAA	2105
Oy	351	AspTrpGlnIyIyValGlnLeuProLeuGlnIyAspGlnValGlyAspLeuGlnIyAserS	371
Db	2106	GCAGCTCAGAGGCTCTCAAAAGATCTTTGAAGGTCAACAAAGGCGACCTTTGAA-----	2157
Oy	371	rGlnGlnIyLyCysAlaProLeuAlaThrGlnIyAlaPheAspGlnIySmetGlnAlaHisGln	391
Db	2158	-GAGAGAGAG-----CGAGGCTCAGATGCTCGGAAGAGCAGCAGCTGTATCTC	2210
Oy	391	nGlnIyAlaIyAlaGlnIyAl-----HisValSerThrValGlnIySth	405
Db	2211	TGAGCTGAAGGAGAGAGACCCGAAAGCCTGGTGGAGCAGCATTAAGCGGAGAAAGAGACT	2270
Oy	405	rGlnGlnGlnGlnGlnIyGlyIyGlyIyGlnIyAlaGlnIyGlyValIyAlaGlnIyIyThrGln	425
Db	2271	GGAAGAAAGAGAGGCTGGGCGCAAGGAGGCTGAGAGGCTCATTAAGCAGCTTGGGGAGGC	2330
Oy	425	yGluSerLeuProProGlu-----LysIleuAlaGlnProGlnGlnIyAlaProGln	441
Db	2331	CCATCAGGCTGAGACTGAATCTCTGCGCGGAGAGCTGGGAGAGGCGCATAGCTGCCACGA	2390
Oy	441	nGlnIyAlaGlnIyProAlaGlnIyGlnIyLeuIySserArgIyIyMetCysValSerGlyGln	460
Db	2391	CACAGCTGAGAGTGAAGTGAAGCAGCTCTGCAAAAGAAAGTGAAGCTGCTGGCGTGAAGGGA	2450
Oy	460	yAspHisThrGlnIyLeuThrAspLeuSerProGlnGlnIyIySthIyLeuProIyHisProGln	480
Db	2451	TGAGGATAGCCAG-----CAAGAGAGAGC-ACAGTATGGCGCCATGTCC	2494
Oy	480	uGlyIyIyLeValSerGlnIyAlaGlnIyMetLeuSerGlnIyAlaGlyIyLeuValGlnGlnIyS	500
Db	2495	AGGAACAGCTGATACCTTTAA-----	2516
Oy	500	rProLeuIyIyLeuIyLeuIySerSerIyGlnIyLeuIySlyIyLeuSerGlnIyIySglnIy	520
Db	2517	-----GAGCAATGTGAGA	2530


```

Db 4474 GCTGAGGTGACCGAGACACAGACACTGCCCCGAGCTGAGTGTG-----ATGACT 4527
Qy 1105 ALaasPvalAspHisValAlaThrCysGlnValIleLysLeuGlnGlnLeuMetGlnGln 1124
Db 4528 GCCAAGTATGAGGGTGC-----MAAGTCAAGGTCTTGAGAGAGAG 4569
Qy 1125 ALaValAlaProGluSerSerGlnThrLeuThrAspSerGlnThrAsnGlySerThrPro 1144
Db 4570 CAGCGGTTCCAGGAAGAGAGAGAACTCACTGCGCCAGGTGAAAGAACGTGATAGAAA 4629
Qy 1145 LeuAlaAspSerAspThrAlaAspGlyThrGlnGlnAsp-----GlnThrIleAspSer 1162
Db 4630 CTGGCTGACTCTGACCAACCCAGCAAGTGCAGACAGACAGACTGAGAGCTGCCAGCT 4689
Qy 1163 Gln-----AspSerLysAlaThrAlaAlaValArgGlnSerGlnValThrGlnGlnGln 1180
Db 4690 CAGGAGGCGAAGAGCAGAGAGAGCCAGCGCTTCAGGCCAGCTGAATGAACTGCA 4749
Qy 1181 AlaAlaThrAlaGlnLysGlnGlnProSer-----ThrLeu 1192
Db 4750 GCCCAGTTGAGCCAGAGAGAGCAGCAGCTGAGCACTATAGCTGCAGATGAGAAAGCC 4809
Qy 1193 ProAsnAsnValProAlaGlnGlnGlnIleHisGlyGlnGln-----ProGly 1207
Db 4810 AAAACACATTATGATGTCGCAAGAGAGCAGCAAAACCAAGACTGCAGAGAGAGCTCCGAGC 4869
Qy 1208 ArgAspValLeuGlnProThrGlnGlnGlnLeuThrAlaAlaValProValLeuAla 1227
Db 4870 CTGAGAGCAGACTCAGAGAGAAACAAAGAGCTGCCAGCTGAAAGCT----- 4914
Qy 1228 LysThrGlnValGlyGlnGlnGlnValAspTrpLeuAspGlyGlnLysValLysGln 1247
Db 4915 ---GAACGCTGGCTGCATGACTACAG-----CAGCTGGGCTGAAGACCAAGAGAG 4962
Qy 1248 GlnGlnGlnValPheValHis-----SerGlyProAsnSerGlnLysAlaAlaAspVal 1265
Db 4963 GCTGAACAGACTGCCCGCACCTTACTGCTCCAGGTCGAGGCTGAGAGCAGAGCTGGCC 5022
Qy 1266 ThrTrpAspSerGlnValMetGlnValAlaGlyCysGln-----GlnLysGln 1281
Db 5023 CATGCAGACACAGACTGTCAGACCTGGCAATTCAGAGTGGCAACTGATGCTTAAAG 5082
Qy 1282 SerThrGlnValGlnSer-----LeuSer 1289
Db 5083 ACCCGGAGACCCAGGCTGAAGCCCACTGATGATGACAGCTGATGATGAC 5142
Qy 1290 LeuGlnGlnGlnGlnMetGlnThrAspValGlnLysGlnLysArgGlnThrLysProGln 1309
Db 5143 TGCAGAGAGGAGACCCCACTCACTATCACAGCAGAGCTGCTGT---ACCCAGCCAGAC 5199
Qy 1310 GlnValSerGlnGlnGlnGlnGlnGln-----ThrAlaAlaProGlnLysGlnGly 1326
Db 5200 GGCACACAGGTCCTCTGGAACACAGCTCACTATCTCCA-GCGCTGCCGCCCAAGGT 5258
Qy 1327 ThrTrpGlyLysProValIleThrLeuAspMetProSerSerGlnLysArgGlyLysAlaLeu 1346
Db 5259 AGAATCCCTGAGAGAGTCTCACTTCACTCCATCCCTGCTCGGAGTCAAGCCCCCT--- 5315
Qy 1347 GlySerLeuGlnGlySerProSerLeuPro----- 1356
Db 5316 ---GAGAGCAGCTGAGCTCCCTGGGAGACGTCCTTCTGCACTCGGTCGTAA 5366
Qy 1357 -----AspGlnAspLys 1360
Db 5367 GACCCGCTCCGCTCGTGGCGCACACGAGATCATCATCATCATCATCATCATCATCATCAT 5423
Qy 1361 AlaGlyCysIleGlnValGlnValGlnSerLeu-AspThrThrValThrGlnThrAlaGln 1380
Db 5424 GCTAGATGTGAAGAGACGACAGCCCACTCATCTCTACAGACCGCGGTCTCTCC 5483
Qy 1380 uAlaValGlnLysValIle-----GlnThrValValIleSerGlnThrGlyGln 1396

```

```

Db 5484 TGCTTCCAGGCTAGCCTGCGAGGACCACTCTTACTACATGCTCTAGCTGCGCTGGT-- 5541
Qy 1396 userProGlnCysValGlyAlaHisLeu----- 1406
Db 5542 -TCTCCCATTTATGCAACATCAGCCCTGCTCAGCTTGCCTGCTACCGCCCCACCACTCG 5600
Qy 1407 ----ProAlaGlnLysSerSerAlaThrGlyGlyHisIleThrLeuGlnHisAlaGlnAs 1425
Db 5601 CAGTTCTCTCTGTGTCTTCCAGGCTCGG-GGTGTCCAGTGGGGCCCTTCAGAAAGAAACA 5659
Qy 1425 PThrValPro-----LeuGlyProGln 1433
Db 5660 GCTTCTACATGGGCACTTGCAGAGATGAGCTGAGCAGCTGATGATGATGAGAAACGCAATTG 5719
Qy 1433 GlnAlaGlnSerIleProIleIleValThrProAlaProGlnSerThrLeuHisProAs 1453
Db 5720 CAGACTCTCAGCAGGCAATGAGTGTGCCCCCACTATGAAAGCTGCTATCTCCCTCG 5779
Qy 1453 pLeuGlnGlnGlnLysSerAlaSerGlnArgLys---SerGlnGlnLysAspLysPr 1472
Db 5780 AGTCCAGGCTTCTCTGAGCTGGGACCATCATCAGATGAGAGAGTGAATAAAGTGAAGCC 5839
Qy 1472 QAsp-----AlaGlyProAspAlaAspGly 1480
Db 5840 CCCAAGAGACCTTGCCTGCGAGCCAGCATGAGCCATCATGATGAGCGAGG 5891

RESULT 13
US-08-931-999-4
; Sequence 4, Application US/08931999
; Patent No. 6043219
;
GENERAL INFORMATION:
APPLICANT: Iandolo, John J.
APPLICANT: Crupier, Scott S.
TITLE OF INVENTION: Broad Spectrum Chemotherapeutic Peptide
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSER: Hovey, Williams, Timmons & Collins
STREET: 2405 Grand Boulevard, Suite 400
CITY: Kansas City
STATE: Missouri
COUNTRY: U.S.A.
ZIP: 64108
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/931,999
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/710,561
FILING DATE: 19-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Collins, John M.
REGISTRATION NUMBER: 26,262
REFERENCE/DOCKET NUMBER: 25043-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 816/474-9050
TELEFAX: 816/474-9057
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 6755 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Staphylococcus aureus
STRAIN: UT0007

```

US-08-931-999-4

Alignment Scores:

Pred. NO.:	4.46e-19	Length:	6755
Score:	410.50	Matches:	319
Percent Similarity:	36.1%	Conservative:	218
Best Local Similarity:	21.48%	Mismatches:	643
Query Match:	5.08%	Indels:	312
DB:	3	Gaps:	55

US-09-902-432-4 (1-1596) x US-08-931-999-4 (1-6755)

169	Qy	LysAspLysAsnGluLysSerAspThrValGlnLeuLeu-----ThrValLysLysAsp	186
15	Db	AGAGAAGAAAAACGAAAGAGAGAGAAAGACAGACAGCAAGAAACAGAGAAAAAGCAG	74
187	Qy	GluGlyGluGlyAlaGluAlaSerValGlyAlaGlyAspHisGlnGluProSerValGlu	206
75	Db	CGGAACACCACCGGGGAGCGCCCGGGGGAAGAACCAACCAACAGCCAACAACGGCG	134
207	Qy	ThrAlaValGlyGluSerAlaSer-----LysGluSerGluLeuLysGlnSerThr	223
135	Db	ACACGCGCGCAGGAAGGAGAGCAAGAGGACACAGAAAGAAAGAACCAAGAACAAAAACA	194
224	Qy	GluLysGlnGluGlyThrLeuLysGlnGluSerSerThrGluLeuProLeuGlnAla	243
195	Db	GAAGAAGAACCAACCA-----AAAGAAAGAAACCAACCCAGAA-----	233
244	Qy	GluSerAspGlnAlaAlaGluGluAlaLysAspGluGlyLysGlnGluLys	263
234	Db	-----AAAGAAGAAAGCAAGAGAAAGGAAGGAAGAAC	263
264	Qy	GluProThrLysSerProGluSerProSerSerProValAsnSerGluThrThrSerSer	283
264	Db	CACAA-AGCGAGAAAAAGCGAAGCGCGAGACAACCGGAACAACGAAAAAACAACAACCA	322
284	Qy	PheLysLysPheThrHisGlyTyrPalaGlyTyrArgLysLysThrSerPheLysLys	303
323	Db	CGCAAAAAAAGCAGCGCAAAACGCAAGCCGCA-----GAAAGAAGCGCAAGAAAAA	376
304	Qy	SerLysGluAspAspLeuGluThrAlaGluLysValArgLysGluGlnGlu-----	319
377	Db	AAAAAAGAAAGACACACAGAAAGACACACAAAAAGAAAGAAAAAACAAGGACACAAACGAA	436
320	Qy	-----AlaGlyLysValAspGluGluGlyLysGluLysThrGluPro	333
437	Db	AGCGGGGGCGACGAAAAAAGAAAGACCAAAACGAGNAAAACACGAAAAAACAACAACCG	496
334	Qy	AlaSerGluGlnGluProAlaGluAspThrAspGlnAlaArgLeuSerAlaAspTyr	353
497	Db	AAGAAAGAAAGAAAGAAACCA-----	517
354	Qy	GluLysValGluLeuProLeuGluAspGlnValGlyAspLeuGluAlaSerSerGluGlu	373
518	Db	---AAGCCCAAGAACCA-----AAAGGGAACGC	544
374	Qy	LysCysAlaProLeuAlaThrGluValPheAspGluLysMetGluAlaHisGlnGluVal	393
545	Db	AAAACGGGCA-----GAAAGAAAAAAGCAAAAGCCAAACCCACACAG	586
394	Qy	ValAlaGluValHisValSerThrValGluLysThrGluGlu-GluGlnGlyGlyGlyG	413
587	Db	GAGAAGGAACAAAAAAGAAAGACAGACAAACACAAAGGAGAGAGGGGAGCCACCA	646
413	Qy	yGluAlaGluGlyValValValGluGly-----	423
647	Db	CGAAAAAAGAAAGCAAAAAAAGAAAGACCGCGAAAGCGCAAGAAAAACGAAGCAACAA	706
424	Qy	-ThrGlyGluSerLeuProProGluLysLeuAlaGluProGlnGluValProGlnGluAl	443
707	Db	AAACGGCAAAACGGGCCCAACAGAAAAA-----CACACCAAAAAAAGCGACAGAAAAA	760
443	Qy	agluProAlaGluLeuMetLysSerArgGluMetCysValSerGlyGlyAspHisThr	463

Db 1771 AAAAAACAAAGCAGAAAAAGAAAGAAACAAACAAACAAACAAAGCACA 1770
 QY rValGluaspSerGluProValGluIleasnGluaspProAsnValProAlaValVa 830
 Db 1771 AAGAAAAAGCGC-----GAGACAAACAAACAAACGCGCGCAAGAC 1812
 QY 1ProLeuSerGluTyrAsnAlaValGluArgGluysMetGluAlaGlnGlyAsnThrGl 850
 Db 1813 CGGACACACGCAAA---CGAAGAAAAAGCGGAACAAAGCAAGAAAGAAAGAAACA 1869
 QY uLeuProGlnLeuLeuGlyAlaValTyrValSerGluGluLeuSerTyrThrLeuValHi 870
 Db 1870 AAAAAACACA-----GCCGACCGCGAGAGAGG----- 1897
 QY sThrValSerValAlaValIleaspGlyThrArgAlaValThrSerValGluGluArgSe 890
 Db 1898 -----GAGACGAAAGAAAGCAAGCAAGCAAGAAAGAAACGCAAA 1935
 QY rProSerTrpIleSerAlaSerValThrGluProLeuGluHiSThr-----AlaGlyGl 908
 Db 1936 A-----GAAAGCAAAACACGAGCCCGCAAGAAAGAAACAAAGAGAGAGAGA 1983
 QY uAlaMetProProValGluGluValThrGluysAspIleIleAlaGly---GluThrPr 927
 Db 1984 AACACAGCCCGAGAAAAAGACAGAAAAACCGAGAAACAAAGAAAGCAGAGAGAGCC 2043
 QY 927 oValLeuThrGlnThrLeuProGluGlyLysAspAlaHisAspAspMetValThrSerGl 947
 Db 2044 CCCAACGAGGAAAGCAAAAGAAACAAACAAACAAAGAGCCACCAACAAAGCAACAGA 2103
 QY 947 uValAspMetThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeuArgThrGl 967
 Db 2104 AGCGAGGGGGAAGAAAGCAAGAAAGAGGACCCAGAACGAAACAAAGACCAAGAGA 2163
 QY 967 uGluValThrGluAlaSerGly-----AlaGl 976
 Db 2164 GAGGAGGCAAAACAGAAAAAGGGGAGAAACAAAGAAAGAAAGAAAGAAAGCCCA 2223
 QY 976 uGluThrThrAspMetValSerAlaValSerGluLeuThrAspSerProAspThrThrGl 996
 Db 2224 AGCAACCGAAAGCCGGAAGGCGCA-----ACGAGAGCAAGAAAGAAACGAG 2271
 QY 996 uGluAlaThrProValGlnGluValGluSerGlyValLeuAspThrGluGluGluArg 1016
 Db 2272 AAGACCAACAGAGG-AGAAAAACAAAGACA-----GAAACCAACAGAGAGCG 2318
 QY 1016 gGlnThrGlnAlaIleLeuGlnAlaValAlaAspLysValLysGluGluSerGlnValPr 1036
 Db 2319 GAAGACCCAAAGAAACAAACAAAGCAACGACAAACAAACAAACAAACAAACAAAC 2361
 QY 1036 oAlaThrGlnThrValGlnArgThrGlySerLysAlaLeuGluLysValGluGluValGl 1056
 Db 2362 ---ACAAGGCGCAACAAAGAAAGAAAGAAAGCAAGAGGAG----- 2403
 QY 1056 uGluAspSerGluValLeuAlaSerGluLysGluLysAspValMetProLysGlyProVa 1076
 Db 2404 -CAAAACAAGCAAGCAAGCCCAACAAACAAACAAACAAACAAACAAACAAACAAAG 2461
 QY 1076 lGln-----GluAlaGlyAlaGluLysLeuAlaGlnGlySerGluThrGluGlnAlaTh 1094
 Db 2462 ACACACCAAGAAAGCGAGCGAGAGAA---CGACACAAACAAAGAAAGAAAGAGAGGGC 2518
 QY 1094 rProGluSerLeuGluValProGluValThrAlaAspValAspHisValAlaThrCysGly 1114
 Db 2519 AAACGAAACCCCGCAAAACCAAGCGCCCGCAAGAGAAACAAAGAAACCAACCAACCA 2578
 QY 1114 nValIleLysLeuGlnGlnLeuMetGluGlnAlaValAlaProGluSerSerGluThrIle 1134
 Db 2579 A-----GAAAGGCGCCACAGAGCGCACAGAGCGCACAGAGAGGCGCAAG 2614
 QY 1134 uThrAspSerGluThrAsnGlySerThrProLeuAlaAspSerAspThrAlaAspGlyTh 1154
 Db 2615 GAGAAACCAACAGAAACAAACAAAGCAAGCAAGCAACAAACCAACAAAGCAGAGAG 2674

QY 1154 rGlnGlnAspGluThrIleAspSerGlnAspSerLysAlaThrAlaValArg----- 1172
 Db 2675 AACAGAGGAAGAGGCCAAAGACCAACAGACAGACAAACAAAGAAAGAGGGAACAAAGACA 2734
 QY 1173 -----GlnSerGlnValThrGluGluGluAlaAlaThrAl 1184
 Db 2735 CAAGAGAGGCAACAGACAAACAAACAAAGCAAGAAAGAAAGAAAGAAAGAGGGGGC 2794
 QY 1184 gLysGluGluProSerThrLeuProAsnAsnValProAlaGlnGluGluHiSGlyGl 1204
 Db 2795 AAAGAACCGCCAAACGCCACGCAAAACAAACAAAGGGGAGAACAGACAGAGCGAGC 2854
 QY 1204 uGluProGlyArgAspValLeuGluProThrGlnGlnGluLeuThrAlaAlaValPr 1224
 Db 2855 AGACGCGCGCAGGAC-----AAACCAACAAACACAGAGGCGAGCGGCAAC 2899
 QY 1224 oValLeuAlaLysThrGluValGlyGlnGluGlyGluValAspTrpLeuAspGlyGluLy 1244
 Db 2900 GAAAGAAACAGAAAGAAACCCGAGAGAGACGAAAGCAAGCAACAAAGCAGAGAGAG 2959
 QY 1244 sValLysGluGluGlnGluValAlaPheValHisSerGlyProAsnSerGlnLysAlaIAs 1264
 Db 2960 AGAAAGAAAGACACAGACACACAGAAAGCCACGCAAAACAAACAAAGAAAGAGGGGA 3019
 QY 1264 pValThrTyrAspSerGluValMetGlyValAlaGlyCysGlnGluLysGluSerThrGl 1284
 Db 3020 CACAAAGCGGGCGAGC-----GGGACAGAAACCAACGAAACCAAG 3061
 QY 1284 uValGlnSerLeuSerLeuGluGluGlyGluMetGluThrAspValGluLysGluLysAr 1304
 Db 3062 ACCGAAACCCAGGCGAGCAAGAGGGGAGAGAGA-AAACAAACAAACAGAAAGACAGAA 3120
 QY 1304 gGluThrLysProGluGlnValSerGluGluGlyGluGlnGluThrAlaAlaProGluHi 1324
 Db 3121 ACA-----CCACAGAAAGCAACCCAAAGAACAGAGAAAGAGGGGGAACCAAGAA 3174
 QY 1324 sGluGlyThrTyrGlyLysProValLeuThrLeuAspMetProSerSerGluArg--Gl 1343
 Db 3175 ACAGAAACCAACCGCCCA-----CAGCCAAACAAAGCCCGCAGGG 3216
 QY 1343 LysAla-----LeuGlySerLeuGlyGlySerProSerLeuProAspGlnAspLysAl 1361
 Db 3217 AAGACAGAGCAAGAGGGGCAAGCAGAGGCGAAC-----GACAAA-- 3256
 QY 1361 aGlyCysIleGluValGlnValGlnSerLeuAspThrThrValThrGlnThrAlaGluAl 1381
 Db 3256 ----- 3256
 QY 1381 aValGluLysValIleGluThrValValIleSerGluThrGluGluSerProGluCysVa 1401
 Db 3257 -----GAAACAAAGCAAGCCGCA----- 3277
 QY 1401 lGlyAlaHisLeuLeuProAlaGluLysSerSerAlaThrGlyGlnIstPrThrLeuGl 1421
 Db 3278 -----CCAAACAAAGAGAGAGCC-----CACGA-AAGCGCA 3311
 QY 1421 nHisAlaGluAspThrValProLeuGlyProGluSerGlnAlaGluSerIleProIleGl 1441
 Db 3312 GCACAGAGAGCAACAGAGAAAGCAACCAACAAAGAGCCACGAGCAC----- 3360
 QY 1441 eValThrProAlaProGluSerThrLeuHisProAsp-----LeuGlnGlyGluIleSe 1459
 Db 3361 -----GAAAGCAAGAAAGAGAGGAGCAGAGAAAGAGAGGCGCGCAACA 3404
 QY 1459 rAlaSerGlnArgGluLysSerGluGluGluAspLysProAspAlaGlyProAspAlaAs 1479
 Db 3405 GAAGAAAGAAAGAGAAAGAAACAAAGCGAAAGAAAGAGAGGAGAGAGAAAGAAAGGA 3464
 QY 1479 pGlyLysGluSerThrAlaIleGluLysValLeuLysAlaGluProGluIleLeuGluLe 1499
 Db 3465 GAACAAACCAACCGAGCAGAGAGAAAGAAAGAAAGGAAAGAAAGAAAGAAAGAAAC 3524

```

Qy 1499 uGluSerLysSerAsnLysIleValLeuAsnValIleGlnThrAlaValAspGlnPheAl 15151
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3525 AGRAACAACAGCAGGAGGAGAAAAAAGAAAAAAGAAACAAG-CAGACAACGAAAGAAACGAGCGGA 3583
Qy 1519 aATgThrGluThrAlaProGluThrHisAlaTyRAspSerGlnThrGlnValProAlaCy 1539
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3584 AAACACGAAAAACGAGAACAGAGAGAAAAAAGAAAAAAGAAACCCCAAGCAGCAACCGCAAGCGAC 3643
Qy 1539 sArgLeuAspSerArgGluProAsnArgCysTrpThrLysMetLysAspAlaLysMetLy 1559
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3644 CAAAAAAGAAAAACGCGGAAACCAACAGG-----AAAAAACCAAAAGAGAGAAAAA 3697
Qy 1559 sHisProValProGlnProArgGluAspLeuGlnValLeuThrValLeuGluAlaTrpAl 1579
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3698 AAAAAGAACGAGCAAGAACGCGAGAAAAACGGAG-----GAAACAAGAAA 3742
Qy 1579 aGlnProArgLys 1583
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3743 AAAACCAAAAAA 3755

RESULT 14
US-07-853-913-1
; Sequence 1, Application US/07853913
; Patent No. 5338839
; GENERAL INFORMATION:
; APPLICANT: McKay, Ronald D.G.
; TITLE OF INVENTION: Nestin Expression As An Indicator of
; TYPE OF INVENTION: Neuroepithelial Tumors
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/853,913
; FILING DATE: 19920319
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/660,412
; FILING DATE: 22-FEB-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/603,803
; FILING DATE: 25-OCT-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/201,762
; FILING DATE: 02-JUN-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/180,548
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: MIT-4641AAAA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11236 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-07-853-913-1

```


QY	283	SerPhelnyslvsPhePheThrhiSGlyTPRlaaglyTPRdxyLvslySerPhe	302
Db	7621	TCCTGCGAAGAA-----GAGAACTCAGAGCCACTGAGGATGTAA	7659
QY	303	LysSerLvsGluAspLeuGluThrAlaGluLvsArgLvsGluAlaGluLvs	322
Db	7650	GAGAGAGAGCCAGATGCTTGAGAGCTGATGAAAAAGAGTCAAGAGTCCCTGAG	7719
QY	323	ValAspGluGluGluLvsGluLvsThrGluProAlaSerGluGluGluLvs	339
Db	7720	TCTCCAGAGAAACCAAGAGATTGGGAACCTCTAGAAAGAGAGATCAAGAAATCTCTG	7779
QY	340	-----ProAlaGluAspThrAspGluAlaArg	348
Db	7780	AGGTATCTTGAGAAACCAAGAGACTTTTGTCACCTAGAAAGCAGAGACAGAGCCCA	7839
QY	349	LeuSerAla-----AspTyrGluLvsValGluLeuProLeuGlu	361
Db	7840	CTGAGATCTCTAGAGTAGAAGAGAGAGAGACAGAAATGTGTAAACTCTAGAAAAAGTG	7899
QY	362	-----AspGluLvsGluAspLeuGluAlaSerSerGluGluLvsCysAlaProLeuAla	379
Db	7900	AGTCAGATTCCTCCGATCTCTA-----GCAAGAGAGAAATGTGAGCAGCATGAGG	7950
QY	380	ThrGluValPheAspGluLvsMetGluAlaHisGluGluValAlaAlaGluValHisVal	399
Db	7951	-----TATCTGAGAAAGAGATCGATCAATATAGAGCTTCTAGAA-----	7992
QY	400	SerThrValGluLvsThrGluGluGluGluGlyGlyGlyGluAlaGluGlyVal	419
Db	7993	-----GACAAAGCTCACAAAGTCTTGGGGTCTCTTGAAAGATAGAAATGGGATAGC	8043
QY	420	ValVal-----GluGlyThrGlyGluSerLeu-----ProProGluLvs	432
Db	8044	ATTATTATACCAAGAAAGTAGAGAGCCAGAGTTCATTAGAGCTCTCAGAAAGAGAGAC	8103
QY	433	-----LeuAlaGluProGluGluValProGluAlaGluPro	445
Db	8104	CAGAGATTTGAAACCATCTAGAAAAAGAAAGTCAAGAGTCTCAGAGCTTCAAGAAAGAA	8163
QY	446	AlaGluGluLeuMetLvsSerArgGluMetCysValSerGlyLysPheHisThrGluLeu	465
Db	8164	GAAAGCGAGGTGATG-----GAGAGATCTCTAGAAAGAGAGAAACCATCAATCACTG	8214
QY	466	ThrAspLeuSerProGlu-----GluLysThrLeuProLysIleProGluLys	481
Db	8215	AGTTTGTAGAAAAAGAGAGACCAATGTTGAGAGCCAACTAGAGAAAGAGTCAAGAC	8274
QY	482	IleValSerGluValGluMetLeuSerSerGluAlaArgIleLysValGluGlySerPro	501
Db	8275	TCAGGGAAGTCTTTGAAGAT-----GAGAGCCAGAGACCTTT-----GGAACCT	8319
QY	502	LeuLysLvsLeuPheSerSerSerGlyLeuLysLysLeuSerGlyLysLysGluLysGly	521
Db	8320	CTGGAAAAA-----GAGATGCGAGATCCCTCAGATCTCTAGCAGAGACAGACCAA-----	8370
QY	522	LysArgGlyGlyGlyAspGluGluProGlyGluTyrGluHisIleHisThrGluSer	541
Db	8371	-----GAGAAACAGAGCTTGAACAAAGAGAC	8397
QY	542	ProGluSerAlaAspGluGluLysGlyGluSerSerAlaSerSerProGluLvs	559
Db	8398	CACCAACACGAGGCGCTGTAGGGAATGACACAGATGCAAGTACCCACCAAAAAAGTG	8457
QY	560	-----ProGluGluThrCysLeuGluLysGlyProLeuGluAlaProGluAspGlyGly	578
Db	8458	GATCCAGAG-----TTACCAAG-----CTCTT-----GGAAT	8487
QY	579	AlaGluGluGlyThrThrSerAspGlyGluLysLysArgGluGlyLysThrProThrAla	598
Db	8488	GACCAAGAAATAGCTAGATCTCTTGAAAAAGAAATCAAGATCAAGTACTTATGT-----	8538

QY	559	SerHelyslySmetValThrProlySlyArgValArgArgProSerGluSerAspLys	618
Db	8539	-----TCACGTGAA	8547
QY	619	GluGluGluLeuGluLysValIlySerAlaThrLeuSerSerThrAspSerThrValSer	638
Db	8548	GA AAAAGTATAGACAGACAGGAACTCTTAA GA-----ACA	8583
QY	639	GluMetGluAspGluValIyStrValGlyGluGluGluLysProGluGluProLysArg	658
Db	8584	GAGATCATAGAACACCTGAGACTGCAGAAAGAGACCTG-----GA AAAGAAAGAG	8634
QY	659	ArgValAspThrSerValSer---TPGValAlaLeuIleCysValGlySerSerIlySly	677
Db	8635	TCCTAGTACTCAGAGCCATTGTG-----TCTACTGAA	8670
QY	678	ArgAlaArgValAspSerSerAspAspGluGlyProArgThrLeuGlyGly---	696
Db	8671	GTGGCTTAGAGACAGTAAACCTCCAGAAAGATGACCCCCAGAGTCCGTAGGCTCTGTG	8730
QY	697	AspSerHisArgAlaGluValAspIlyAspLysAspLysGluValaGlyThrAspAlaValPro	716
Db	8731	GATGAGAAACGAGAGACACTGACATCCCTTGA AAAAGAAAGTCA-AGAAGTACGCTCTC	8789
QY	717	AlaSerThrGluGluGluAspGluAlaGluGlySerSerSerProGluProAlaGlySer	736
Db	8790	GGGCAAGTGAACCTTAGAGCCAGGGTAAAGGACAGTCCGCA-----GTGCTGCAGATGAA	8846
QY	737	ProSerGluGlyGluGlyValaSerThrTPGLeuSerPheLysArgLeu-----ValThr	754
Db	8847	AGAGGCTCTGACAGAGAAC-AGCACCA-GAGCTCTAGAGAGAGTGAACAGAGACTG	8904
QY	755	ProArgLysLysSerIlySerLysLeuGluGluValaGluAspSerSerValGluGlu	774
Db	8905	CCTAGCTGTGAATCAACAGCGGTGGAGAGATGTGTGAGAGGCAAGCAGTGGTCTAG	8964
QY	775	-----LeuSerThrGluIleGluProSerArgGluGluSerThrValSerIlySly	791
Db	8965	GAAGACCTCTGGCCACACAGAGATGGAGATGAGATTAAGCAGAGTGCATCTG---	9021
QY	792	LysPheIleProGlyArgGlySlyArgAlaAspGlyLysGluGluAlaThrVal	811
Db	9022	-----AGGGGCAAGGTGAGAGAAAGACGTGCACCA	9054
QY	812	GluAspSerGlyProValGluIleAsnGluAsp-----	822
Db	9055	GAGGGA-----GAGCTGTTCAGAGATATTGTGGGGAGAGCCGTGAGTCTGGGG	9102
QY	823	-----AspProAsnValProAla-----	828
Db	9103	AGCTCTGAGCCCAAGGAGAGAGGTCCTCTGTAGAGCCCTCGACACCTGAGAGAGAGG	9162
QY	829	-----ValValProLeuSerGlyIyr-----AsnAlaValGluArgGluLysMetGlu	844
Db	9163	GCCTTAGAGGTCACAGTGTCTAGTCAATGCCAGAGCTACACAGGCGAGATGAGATTGA	9222
QY	845	AlaGluGlyAsnThrGluLeuProGluLeuGlyAlaValIyrValSerGluGluLeu	864
Db	9223	GCCCAAGAGGTGAACA-----GACTCCATA	9249
QY	865	SerLysThrLeuValHisThrValSerValAlaValIleAspGlyThrArgAlaValThr	884
Db	9250	GAGGTGACCTT-----GGGTGAGAGCTTCGACAACTGAGACTG	9288
QY	885	SerValGluGluArgSerProSerThrIleSerAlaSerValThrGluProLeuGluHis	904
Db	9289	GAACCTCGACAGGA-----GTGGTAGGGCTTAGAGACCCAAAGGCAATTT	9333
QY	905	ThrIlaGlyGluAlaMetProProValGluGluValIyrThrGluLysAspIleIleGlu	924
Db	9334	GCCAGGAGAGGCGCATTTCCCA-----TCCCTGGGGAGAAAGTGAAGCAAG	9387
QY	925	GluThrProValLeuThrGluThrLeu-----ProGluGlyLysAsp	938

```
Db 9388 -----ATAGCTCAGGGCTTGAAGGGCTCGAAGAGCAACAAAGAGGAGGT 9435
      ||| ||| ||| ||| |||
Qy 939 AlaHisaspMetValThrSerGluValaspPheThrSerGluAlaValThrAlaThr 958
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 9436 GCTCTGACTCGGGATCCTGAATTCGCCAAGACTAGCAGGAGGCTCTGGAATGCCAG 9495
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 959 -----GluThrSerGluAlaLeuArgThrGluGluValThrGluAlaSerGlyAlaGlu 976
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 9496 GGCCATGAAGAGTCTGAGTCCATGAGGGCTGGGAAGAGGAGGCTCA-----CTG 9549
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 977 GluThrThrAspMetValSerAlaValSerGlnLeuThrAspSerProAspThrThrGlu 996
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 9550 GAGACTTCAGAT-----CATGAGGCGAGTGATGCCCTTCAGCCAGGCC 9594
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 997 GluAlaThrProValGlnGluValGluSerGlyValLeuAspThrGluGluGluArg 1016
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 9595 CCAGAAACA-----GAGGAAGATGAG 9615
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1017 GlnThrGlnAlaLeuGlnAlaValAlaAspLysValLysGlu----- 1031
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 9616 GGTGCACAGGAGCAGCTGACAGCCCTGGTCCCAAGCTTTGGAACCCCTGTTCACCCATC 9675
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1032 -----GluSerGlnValProAlaThrGlnThrValGln----- 1042
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 9676 CCAATCTGACAGATGCCATGAGCTCAG-----CCCCAGGCTGAGGGGATCCAGAGGCT 9732
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1043 -----ArgThrGlySerLysAlaLeuGluLysValGluGluValGluGluAsp 1058
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 9733 GGCTGGCAGCCAGAGCTGGTCTGAAGCAGCTAGAAAGGCTGAAATGAGCCAGAG--- 9789
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1059 SerGluValLeuAlaSerGluLysGluLysAspValMetProLysGlyProValGlnGlu 1078
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 9790 -----TTTGTCTTGGGGAGATCCCGAG 9813
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1079 AlaGlyAlaGluHisLeuAlaGlnGlySerGluThrGlyGlnAlaThrProGluSerLeu 1098
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 9814 ---GGCTCCAGGATTCGGGAAGAGGGCAGAGAAAGCAGGAGCAGATGATCTAGGGGAA 9870
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1099 GluValProGluValThrAlaAspValAspHisValAlaThrCysGlnValLysLeu 1118
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 9871 ACTCTCCTGACTACTCTCCCTGGGCTCTACTGAGGTCCCTGCTCTCCAAAGTGG 9930
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1119 GlnGlnLeuMetGluGlnAlaValAlaProGluSerSerGluThrLeuThrAspSerGlu 1138
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 9931 GATCTGGCTGGAGAACAGAGGCTTTCCTCAAGGGATGCCGGAAGGAGGAG--- 9984
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1139 ThrAsnGlySerThrProLeuAlaAspSerAspThrAlaAspGlyThrGlnGlnAspGlu 1158
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 9985 TGGGTCTCTGCTGCTCCCTGAGGCTCAGTGGTCCACCGGAGAGGAGGAGGAG 10044
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1159 ThrIleAspSerGlnAspSerLysAlaThrAlaAlaValArgGlnSerGlnValThrGlu 1178
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10045 CAAGGCCATGGCTCTGAC-----CTATCATCTGAGGAGTTT 10080
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1179 GluGluAlaAlaThrAlaGlnLysGluGluProSerThrLeuProAsnAsnValProAla 1198
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10081 GAGGACCTAGGACT-----GAGGCTCTCTCTTCCA---GGGTTCCCAAG 10125
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1199 Gln-----GluGluHisGlyGluGluProGlyArgAspValLeuGluProThrGlnGln 1216
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10126 GAGTGGCAGATCAGCTGGGCCAAGTGCCTC-----CCGGTACTGACGCTGATGCTGG 10179
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1217 GluLeuThrAlaAlaValProValLeuAlaLysThrGluValGlyGlnGluGlyGlu 1236
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10180 GATCAGGGTGGGATCTGATGGTTGCTGATGAGAAAGAAAGTGGGGAGGAGGAGAG 10239
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1237 ValAspTrpLeuAspGlyLysValLysValLysGluGlnGluValPheValHisSerGly 1256
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10240 GAAGAAGATGCTGAT---GAGGAAGGAGCAGAGTCAGGAGCTCAGTGGTGGGTCAGGG 10296
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1257 ProAsnSerGln-----LysAlaAlaAspValThrTyrAsp-----SerGlu 1270
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
Db 10297 GCCTCTGGTGGAGGCTGCAAGGTCAGGATATTGCCCCAAAGAGGAGACCCGGTACAGGAG 10356
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1271 ValMetGlyValAlaGlyCysGlnGluLys-----GluSerThrGluValGln 1286
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10357 TCTTGGGTGTCTAGTGTCTCTGGGATGATGGCTTGAGAGGTCTGCAGCAATGTTCT 10416
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1287 SerLeuSerLeuGluGluGluGluMetGluThrAspValGluLysGluLysArgGluThr 1306
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10417 GCCTTAGAGATG-----GTA 10431
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1307 LysProGlnValSerGluGluGluGluGlnGluThrAlaAlaProGluHisGluGly 1326
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10432 TCTCAGCAGACAGTCTCAGCCTTCTGGTCTCAGAGAGTCTGAGTCTGCTCTCTTGGAGGG 10491
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1327 ThrTyrGlyLysProValLeuThrLeuAspMetProSerSerGluArg----- 1342
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10492 GAGGAAGGTCAAGTGAAGTACTGACCATTTAGATGCTCCCCAGGAGGTGACCATGGTCCCG 10551
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1343 -----GlyLysAlaLeuGlySerLeuGlyGlySerProSerLeuProAspGlnAspLys 1360
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10552 GGGGTAGGAGATCCCTTTCACATTTGGTGGCCAGAGCCCAACTTGGACTCAGAACCAAGTG 10611
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1361 AlaGlyCysIleGluValGlnValGlnSerLeuAspThrThrValThrGlnThrAlaGlu 1380
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10612 AATGGGAATATGGAATGAGATGAGACAGGCTGAGGGGAGGTGCTCTGGATGGGGAC 10671
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1381 AlaValGluLysValIleGluThrValValIleSerGluThrGly----- 1395
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10672 GAGGATCAAGAACTCCTA-----TTACAGGGACAGAGGTGGTGTCTCTAAAGTTTCT 10725
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1396 -----GluSerProGluCysValGlyAlaHisLeuLeuProAlaGluLys----- 1410
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10726 TTGGTAGCATCTCCT-----GTGCATCTAGGCCCAAGCCAGCCCTGAAAGTTC 10773
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1411 -----SerSerAlaThrGlyGlyHisThrThrLeuGlnHisAlaGluAspThrValPro 1428
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10774 ACTCTAGTGGGTAGATGGGATCTCTGTC-CTCAGGGGAAGACTAGAAACTGCCCT 10832
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1429 Leu 1429
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10833 CTG 10835
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
RESULT 15
US-08-056-200-93
; Sequence 93, Application US/08056200
; Patent No. 5616500
; GENERAL INFORMATION:
; APPLICANT: Steinert, Peter M.
; APPLICANT: Lee, Seung-Chul
; APPLICANT: Kim, In-Cyu
; APPLICANT: Chung, Soo-Il
; APPLICANT: Park, Sang-Chul
; TITLE OF INVENTION: Trichohyalin and Transglutaminase-3 and
; TITLE OF INVENTION: Methods of Using Same
; NUMBER OF SEQUENCES: 117
; CORRESPONDENCE ADDRESS:
; ADDRESS: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive, Sixteenth Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/056,200
; FILING DATE: 30-APR-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fedrick, Michael F.
```


QY 533 luTyrGlnHisIleHisThrGluSerProGluSerAlaAspGluGlnLysGlyGluSerS 553
DB 4223 TCAGACGACGACTGAACGGCGAGAGTGGAGAGACTCGAGCAGGAGGAGCGCGACG 4282
QY 553 exAlaSerProGluGluProGluGluThrThrCysLeuGluLysGlyProLeuGluA 573
DB 4283 AGCGGCTGAACGGCGAGGAGCGGAGAGAGAGCGGCCACGAGCTCTGAAGAGCGAGG 4342
QY 573 laProGlnAspGlyGluAlaGluGluThrThrSerAspGlyGluLysArgGluG 593
DB 4343 AGCAGGAGGAGCGCGCCACGAGCACTGAGCGCGAGCAGCAGCAAGAGCGCGAGCAGC 4402
QY 593 lyIleThrProTrpAlaSer-----PheLysLysMetValThrP 606
DB 4403 GGCTGAACGGCGAGGAGGAGAGAGAGGCTCGAGCAGCGCTGAAGCGCGCATGAGG 4462
QY 606 rolyLysArgValArgProSerGluSerAspLysGluGluGluLeuGluLysValL 626
DB 4463 AAGAGAGCGCGAGCAGGAGCTAGCTGAGGAGGAGCAGAGCGCGCGGAGCGGATTA 4522
QY 626 ysSer-----AlaThrLeuSerSerThrAspSerThrValSerG 639
DB 4523 AGAGCGCATCCCGAGTGGCAGTGGCAGCTAGAAAGCGAGGCGCGACCGCGCAAGCA 4582
QY 639 luMetGlnAspGluValLysThrValGlyGluGluGlnLysPro-----GluGluProL 657
DB 4583 AAGTCTTACTCGAGGCGCGCCCAAGCAGGAGGCGGAGGCGCGCCCAAGAGCAGGAGGAAA 4642
QY 657 ysArgArgValAspThrSerValSerTrpGluAlaLeuIleCysValGlySerSerLysL 677
DB 4643 AGAGCGCGCGAGGAGTGCAGTTCGCAATGGCAG-----GAGGAGG 4681
QY 677 ysArgAlaArgLysAlaSerSerSerAspGluGlyGlyProArgThrLeuGlyGlyA 697
DB 4682 AACGGGCTCACCGCGCAGCAGCAGGAGGAGCAGCGCGGAGCTTCACATGG-----4734
QY 697 spSerHisArgAlaGluGluAlaSerLysAspLysGluAlaGlyThrAspAlaValProA 717
DB 4735 --CAGTGGCGCGCGGAGGAGAAAGAGC-----GAGAGGGCGCGTCAGAGGCTGTCCG 4783
QY 717 laSerThrGlnGlnAspGln-AlaGlnClySerSerSerProGluProAlaGlySer 736
DB 4784 CAGGCGCCCATTCGGGGAGCAGCGGAGAGGCGAGCTGAGGGCGCGAGGCGCGCAGCAGC 4843
QY 737 ProSerGluGly-----740
DB 4844 GGGACAAACGGTTCTCCCGGAGGAGGAGAGAGGAGCAGCGCGCGCGCGCGCAGCGCG 4903
QY 741 -----GluGlyValSerThrTrpGluSerPheLysArgLeuValThrProArgLys 757
DB 4904 AGAGGGAGAAAGAGCTGCAGTTCCTCGAGG-----AAGAGGAGCAGCTCCAG-CGGCGG 4956
QY 758 LysSerLysSerLysLeuGluGluLysAlaClnAspSerSerValGluGlnLeuSerThr 777
DB 4957 GAGCGTCCCAACAGCTCCAGGAGGAGGAGGAGCGCGCTCCAGGAGGATCAGGAGGAGG 5016
QY 778 GluIleGluProSerArgGluGluSerTrpValSerIleLysLysPheIleProGlyArg 797
DB 5017 CGACAGGAGCAGCGCGCGGACCAAAATGG-----AGGTGGCACTAGAGAGAGAA 5067
QY 798 ArgLysLysArg-----AlaAspGlyLysGlnGluGlnAlaThrVal 811
DB 5068 AGGAAGAGAGCGCGCCACACGCTGACGCCAAGCAGCGCCCTACAAGAGCAGCTCAGGAG 5127
QY 812 GluAspSerGlyProValGluIleAsnGluAspAspProAsnValProAlaValPro 831
DB 5128 GAACAGCAG-----CTGCTCAGGAGGAGGAG-----5154
QY 832 LeuSerGluTyrAsnAlaValGluArgGluLysMet-----843
DB 5155 ---GAGGAGCTACAGAGAGGAGCGCGGAGAGAGAGGCGGCCAAGAGACAGGAGAGCAA 5211

QY 844 -----GluAlaGlnGlyAsnThrGluLeuProGlnLeuGlyAlaValTyrVal 860
DB 5212 TACCGCGAGGAGAGCAGCTGCAGCAGGAGAAAGAGCAGCTGCTG-----5256
QY 861 SerGluGluLeuSerLysThrLeuValHisThrValSerValAlaValIleAspGlyThr 880
DB 5257 AGAGGGAACCGGAGAAAGAAG-ACGCCAGGAGCGGAAAGCAATATCGGAAGGATAA 5315
QY 881 ArgAlaValThr-----SerValGluGluArgSerPro 891
DB 5316 GAAGCTCGCAGCAGAGAGCAGCTGCTGGGAGAGAAACCGGAGAGAAAGAGCGCCCA 5375
QY 892 SerTrpIleSerAlaSerValThrGluProLeuGluHisThrAlaGlyGluAlaMetPro 911
DB 5376 GGA-----GCGGAGAAATAATACCG 5396
QY 912 ProValGluGluVal-----ThrGluLysAspIleIleAlaGlu 924
DB 5397 CGAGGAGAGGAGCTTCAGCAGGAGGAGAGCAGCTGCTGAGAGAGGAACGGGAGAGAG 5456
QY 925 GluThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHisAspMetVal 944
DB 5457 AAGCGCGCAGGAGTGGGAGAGGAGTACCGCAAAAGACGAGCT-----5501
QY 945 ThrSerGluValAspPheThrSerGluAlaValThrAlaThr-GluThrSerGluAlaLe 964
DB 5502 -----GCAGCAGGAAGAGCAGCTGCT 5525
QY 964 uArgThrGluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSerAl 984
DB 5526 GAGAGGAAACGGGAGAAAGA-----5547
QY 984 aValSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGluVa 1004
DB 5548 -----AGACTCCAGGAGCGGAGAGGCAATATCGGAGGAGAGAGGAGCTGCAGCAGGA 5600
QY 1004 lGluSerGlyValLeuAspThrGluGluGluArgGlnThrGlnAlaIle-----1021
DB 5601 GGAAGAGCAGCTCTGGGAGAGACGGGAGCAGAGAGCGCAGGAGCTGGAGAGGCA 5660
QY 1022 -LeuGlnAlaValAlaAspLysValLysGluGluSerGlnVal-----ProAl 1037
DB 5661 ATATCGAAGGAGAGAGAGCTGCAGCAGGAGGAGAGAGCAGCTGCTGAGAGAGGAGAACCGGA 5720
QY 1037 aThrGlnThrValGlnArgThrGlySerLysAlaLeuGluLysValGluGluValGluG 1057
DB 5721 GAAGAGAGCGCGCAGGAGCGGAGAGGCAATGTCGGAGGAGAGAGAGCTGCAGCAGGA 5780
QY 1057 uAspSerGluValLeuAlaSerGluLysGluLysAspValMetProLysGlyProValG 1077
DB 5781 GGAAGAGCAGCTGCTGAGAGAGAAACGGGAGAAAGAG-AAGGCGCCA-----5825
QY 1077 nGluAlaGlyAlaGluHisLeuAlaGlnGlySerGluThrGlyGlnAlaThrProGluSe 1097
DB 5826 -----GGAGCTGGAGAGCAATATCGGGAGGAGGAGAGCTTCAGCGCCAGAAAAGGA 5878
QY 1097 rLeuGluValProGluValThrAlaAspValAspHisValAlaThrCysGlnValIleLy 1117
DB 5879 AACAGCGATACCGGATGAGGATCA-----5903
QY 1117 sLeuGlnGlnLeuMetGluGlnAlaValAla-ProGluSerSerGluThrLeuThrAspS 1137
DB 5904 ----GCGCAGTGATCTGAATGCGAGTGGGAACCGAGAAAAGAAATGCAATGCTGCTGATA 5959
QY 1137 er-----GluThrAsnGlySerThrProLeuAlaAspSerAspT 1150
DB 5960 ACAAGGTTTACTCAAGGCGAGAGAGAGTGAACAGTTCGCGCAGTTGGAGAGATTCCAGG 6019
QY 1150 hrAlaAspGlyThrGlnGlnAspGluThrIleAspSerGlnAspSerLysAlaThrAlaA 1170
DB 6020 TGGCGGAGCAGACAAATCCAG-----CAAGATCTGCACACCTGCTGG 6061
QY 1170 laValArgGlnSerGlnValThrGluGluGluAlaAlaThrAlaGlnLysGluProS 1190

```

Db 6062 GTGAACAGGAAAGAGAGATCTGTGACAGAGAGAGAGCGCTGGCAGCGCCAAACAGGC 6121
      ::::: ::::: ||::: ||:::
Qy 1190 eRThLeuProAsnAsnValProAlaGlnGluGlnHisGlyGluGluProGlyArgAspV 1210
      ::::: ||::: ||:::
Db 6122 ATTTCCAGAGAGAAACAGCTGAGAGCGAGAAAGAAAGAGCAAAAGAGCGGAGAA 6181
      ::::: ||::: ||:::
Qy 1210 aLeuGluProThrGlnGlnGluLeuThrAlaAlaValProValLeuAlaValThrG 1230
      ::::: ||::: ||:::
Db 6182 GGAACTCCCAAGAGAGAAAGCAGTTGCTGAGAGAG-----GAAAGAGAG 6226
      ::::: ||::: ||:::
Qy 1230 luValGlyGlnGlnGluValAspTrpLeuAspGlyGluValLysGlnGlnGlnG 1250
      ::::: ||::: ||:::
Db 6227 AGAAGAGAGCGCTCAAGAGAGAGAG-----AGAAATTTCCCGAGAGAGAGAA 6274
      ::::: ||::: ||:::
Qy 1250 luValPheValHisSerGlyProAsnSerGlnLysAlaAlaAspValThrTyraSpSerG 1270
      ::::: ||::: ||:::
Db 6275 AGCTGCTC-----CAGAAAGAGAGAGAGAGCGCGCTCCGCAAG 6316
      ::::: ||::: ||:::
Qy 1270 luValMetGlyValAlaGlyCysGlnGlnLysGluSerThrGluValGlnSerLeuSerL 1290
      ::::: ||::: ||:::
Db 6317 AG-----CGTACAGAGAAATTTCCGAGAGAGAGAGAGAGCTCCATC 6355
      ::::: ||::: ||:::
Qy 1290 euGluGluGly-----GluMetGluThrAspValGluLysGluLysA 1304
      ::::: ||::: ||:::
Db 6356 AGGAAACAAGGGAGAAATTTCTCGAGAGAGAAACAGCGCTCGCGAGAGAGAGAGAA 6415
      ::::: ||::: ||:::
Qy 1304 rgGluThrLysProGluGlnVal-----SerGluGluGlnGlnGlnGlnThrAlaAlaP 1322
      ::::: ||::: ||:::
Db 6416 AATTCCTTAAGAGAGAGAGAGAGAGCTGCGCTCGAGAGAGCGGAGAGAGAGAGAGAG 6474
      ::::: ||::: ||:::
Qy 1322 roGluHisGlu-----GlyThrTyGlyLysProValLeu----- 1333
      ::::: ||::: ||:::
Db 6475 CGCGACAGAGAAATTTCCGAGAGAGAGAGAGAGAGCTGAGCGCGCAAGAGAGAGAGAGAA 6534
      ::::: ||::: ||:::
Qy 1334 -----ThrLeuAspMetProSerSerGluArgGlyLysAlaLeuGlySerL 1349
      ::::: ||::: ||:::
Db 6535 TTCCTGAGAGAGAGAGAGAGAGAGTGCCTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6590
      ::::: ||::: ||:::
Qy 1349 euGlyGlySerProSerLeuProAspGlnAspLysAlaGlyCysLLeGluValGlnValG 1369
      ::::: ||::: ||:::
Db 6591 --GAGAGAGAGAGAGAGAGAGTGCCTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6630
      ::::: ||::: ||:::
Qy 1369 InSerLeuAspThrThrValThr-----GlnThrAlaGluAlaValGluLysValI 1386
      ::::: ||::: ||:::
Db 6631 CCAGAGAGAGAGAGAGAGAGTCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6690
      ::::: ||::: ||:::
Qy 1386 leGluThrValAlaIleSerGluThrGlyGluSerProGluCysValGlyAlaHisLeuL 1406
      ::::: ||::: ||:::
Db 6691 GACAGAGAAATTTCTGAGAGAGAGAGAGAGAGTGCCTGCGC----- 6728
      ::::: ||::: ||:::
Qy 1406 euProAla-----GluLysSerSerAlaThrGlyGlyHisTrpThrLeuGlnHis--A 1423
      ::::: ||::: ||:::
Db 6729 --CCAAGAGCGTACAGAGAAATTTCCGCGAACAAGAG-----ACTGCGAGTCAAGAA 6777
      ::::: ||::: ||:::
Qy 1423 laGluAspThrValProLeuGly-----ProGluSerGlnAlaGluSerI 1438
      ::::: ||::: ||:::
Db 6778 CCAGAGAGAGAAATTTCTCGAGAGAGAGAGAGAGAGTGCACCGCAGAGAGAGAGAGAG 6837
      ::::: ||::: ||:::
Qy 1438 lePro-----IleIleValThrProAlaProGluSerThrLeu-HisProAsp 1453
      ::::: ||::: ||:::
Db 6838 TTCTCTCCAGAGAGAGAGAGAGAGTGCCTCGCGAGAGAGAGAGAGAGAGAGAGAGAG 6897
      ::::: ||::: ||:::
Qy 1454 leuGlnGlyGluIleSerAlaSerGlnArgGluArgSerGluGluLysAspProAsp 1473
      ::::: ||::: ||:::
Db 6898 CGTGACAGAGAAATTTCCGAGAGAGAGAGAGAGTGCCTCGAGAGAGAGAGAGAGAGAGAG 6957
      ::::: ||::: ||:::
Qy 1474 AlaGlyProAspAlaAspGlyLysGluSerThrAlaIleGluLysValLeuLysAlaGlu 1493
      ::::: ||::: ||:::
Db 6958 AGCGCGCAAGAGAGAGAGAGAGAGAAATTTCCGTTAGAGAGAGAGAGAGAGAGAGAGAG 7017
      ::::: ||::: ||:::
Qy 1494 ProGluIleLeuGluLeuGluSerLysSerAsnLysIleValLeuAsnValIleGlnThr 1513
      ::::: ||::: ||:::

```

```

Db 7018 CAAGAGAGAAATTTCAATGAGAGAGAGAG----- 7044
      ::::: ||::: ||:::
Qy 1514 AlaValAspGlnPheAlaArgThrGluThrAlaProGluThrHisAlaTyraSpSerGln 1533
      ::::: ||::: ||:::
Db 7045 -----CAGCAGAGTGCCTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7098
      ::::: ||::: ||:::
Qy 1534 ThrGlnValProAlaCysArgLeuAspSerArgGluProAlaAsnArgCysTrpThrLysMet 1553
      ::::: ||::: ||:::
Db 7099 TTCGCGAGAGAGAGAGAGAGTGCCTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7158
      ::::: ||::: ||:::
Qy 1554 LysAspAlaLysMetLysHisProValProGlnProArgGluAspLeuGlnValLeuThr 1573
      ::::: ||::: ||:::
Db 7159 CGTGACAGAGAAATTTCTCGAGAGAGAGAGAGAGTGCCTCGCGAGAGAGAGAGAGAGAG 7217
      ::::: ||::: ||:::
Qy 1574 ValLeuGluAlaTrpAlaGlnProArgLysCysLeuProArgLeuGlnLeuLysAlaPro 1593
      ::::: ||::: ||:::
Db 7218 -----GCTGCGTCAAGAGAGAGAGAGAGAGAAATTTCCGTGAGAGAGAGAGAGAG 7268
      ::::: ||::: ||:::

```

Search completed: December 13, 2002, 04:17:51
 Job time : 438 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 13, 2002, 03:13:21 ; Search time 944 Seconds
(without alignments)
668.597 Million cell updates/sec

Title: US-09-902-432-4
Perfect score: 8073
Sequence: 1 MNGSGTGRSPEQPAGSDT.....AWAPRKCLPRLQKAPVSK 1596

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 355320 seqs, 197730502 residues

Total number of hits satisfying chosen parameters: 710640

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool/US0902432/runat_12122002_141453_25918/app_query.fasta_1.1735
-DB=Published Applications NA -QFMT=fastap -SUPFI=xrnpb -MINMATCH=0.1
-LOPCLI=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -LOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US0902432.cgn 1.1 85 @runat_12122002_141453_25918
-NCPU=6 -ICPU=3 -NO XLPXY -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	8073	100.0	6160	9 US-09-902-432-3	Sequence 3, Appli
2	6005	74.4	5134	9 US-09-902-432-1	Sequence 1, Appli
3	4300.5	53.3	5346	10 US-09-738-877-2	Sequence 2, Appli
4	4300.5	53.3	6608	10 US-09-738-877-1	Sequence 1, Appli

5	4300.5	53.3	6608	10	US-09-880-107-3439	Sequence 3439, Ap
6	462	5.7	5361	9	US-09-742-096-2	Sequence 2, Appli
7	462	5.7	6152	9	US-09-742-096-1	Sequence 1, Appli
8	431.5	5.3	7997	10	US-09-864-761-17726	Sequence 17726, A
9	427	5.3	15231	10	US-09-917-800A-1505	Sequence 1505, Ap
10	418	5.2	14800	10	US-09-954-456-1501	Sequence 1601, Ap
11	411	5.1	7104	10	US-09-815-242-4580	Sequence 4580, Ap
12	411	5.1	7107	10	US-09-815-242-8291	Sequence 8291, Ap
13	408	5.1	10300	10	US-09-960-253-145	Sequence 145, App
14	398.5	4.9	4840	10	US-09-880-107-3423	Sequence 3423, Ap
15	395	4.9	11185	10	US-09-880-107-3311	Sequence 3311, Ap
16	388	4.8	3953	10	US-09-864-761-19041	Sequence 19041, A
17	382.5	4.7	7639	9	US-09-971-536-28	Sequence 28, Appli
18	382.5	4.7	7792	12	US-10-044-090-359	Sequence 359, App
19	379	4.7	7497	10	US-09-960-253-175	Sequence 175, App
20	378.5	4.7	10432	10	US-09-919-172-97	Sequence 97, Appli
21	377	4.7	7434	10	US-09-815-242-4761	Sequence 4761, Ap
22	377	4.7	7437	10	US-09-815-242-8869	Sequence 8869, Ap
23	371	4.6	6940	10	US-09-880-107-1724	Sequence 1724, Ap
24	365.5	4.5	5331	10	US-09-764-176-6	Sequence 6, Appli
25	355	4.4	4854	9	US-09-963-875-2	Sequence 2, Appli
26	351.5	4.4	1855	10	US-09-864-761-19708	Sequence 19708, A
27	346.5	4.3	4372	10	US-09-864-864-323	Sequence 323, App
28	346.5	4.3	4372	10	US-09-864-864-337	Sequence 337, App
29	345	4.3	10825	10	US-09-727-384-5	Sequence 5, Appli
30	342.5	4.2	7065	9	US-09-991-496-115	Sequence 115, App
31	342.5	4.2	7065	10	US-09-874-923-115	Sequence 115, App
32	341	4.2	3388	9	US-09-954-531-988	Sequence 988, App
33	341	4.2	3388	9	US-09-954-531-1382	Sequence 1382, Ap
34	341	4.2	3388	10	US-09-954-456-1602	Sequence 1602, Ap
35	341	4.2	3388	10	US-09-967-768A-245	Sequence 245, App
36	335.5	4.2	4780	10	US-09-962-436-287	Sequence 287, App
37	335.5	4.2	7596	10	US-09-954-456-2215	Sequence 2215, Ap
38	335	4.1	1958	10	US-09-864-761-2927	Sequence 2927, Ap
39	334.5	4.1	7035	10	US-09-815-242-8615	Sequence 8615, Ap
40	329.5	4.1	6114	9	US-09-938-842A-2700	Sequence 2700, Ap
41	327	4.1	6457	10	US-09-880-107-3389	Sequence 3389, Ap
42	323.5	4.0	5721	10	US-09-785-770A-15	Sequence 15, Appli
43	323.5	4.0	8121	10	US-09-785-770A-14	Sequence 14, Appli
44	323	4.0	7869	10	US-09-954-456-1921	Sequence 1921, Ap
45	316	3.9	4747	9	US-09-854-133-729	Sequence 729, App

ALIGNMENTS

RESULT 1
US-09-902-432-3
; Sequence 3, Application US/09902432
; Patent No. US20020160002A1
; GENERAL INFORMATION:
; APPLICANT: Irwin H. Gelman
; APPLICANT: Susan G. Jaken
; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
; FILE REFERENCE: A30558-A-FWC-A 070156.0597
; CURRENT APPLICATION NUMBER: US/09/902,432
; CURRENT FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 08/978,277
; PRIOR FILING DATE: 1997-11-25
; PRIOR APPLICATION NUMBER: 08/665,401
; PRIOR FILING DATE: 1996-06-18
; PRIOR APPLICATION NUMBER: 08/635,121
; PRIOR FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 6160
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-09-902-432-3

Alignment Scores:
Pred. No.: 0
Score: 8073.00
Length: 6160
Matches: 1596

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
US-09-902-432-4 (1-1596) x US-09-902-432-3 (1-6160)
QY 1 MetGIYAlaGlySerSerThrGluGlnArgSerProGluGlnProAlaGlySerAspThr 20
DB 33 ATGGCGGAGGAGGAGTTCACCGAGAGCGAGCCCGAGACACCGCGGGGAGCAGACAG 92
QY 21 ProSerGluLeuValLeuSerGlyHisGlyProAlaAlaGluAlaSerGlyValAlaGly 40
DB 93 CCGAGCGAGCTGCTGCTCAGTGCCATGCGCCAGCTGAGACCTTCGGGAGAGCTGGA 152
QY 41 AspProAlaAspAlaAspProAlaThrLysLeuProGlnLysAsnGlyGlnLeuSerSer 60
DB 153 GACCCCGCGGAGCGGAGCCCGCACCAAGCTCCACAGAAAGATGGCCAGCTGCTTCT 212
QY 61 ValAsnGlyValAlaGluGlnGlyAspValHisValGlnGluGluAsnGlnGlyGln 80
DB 213 GTCAACGGCGTGTGAAACAGAGAGATGTCCATGTCCAAAGAGAAACACAGAGGGGCGAG 272
QY 81 GlnGluGlnValValAspGluAspValGlyGlnArgGluSerGluAspValArgGluLys 100
DB 273 GAGGAAGAAAGTGTGATGATGATGATGAGACAGCGAGGCTCAGAAAGTGTGAGAGAAAAA 332
QY 101 AspArgValGluGluMetAlaAlaAsnSerThrAlaValGluAspL1eThrLysAspGly 120
DB 333 GACCGAGTGAAGAAATGGCGGCACTCCACAGCTGTGGAGATATCAAAAGATGG 392
QY 121 GlnGluGluThrSerGluLeileileileGlnL1eProAlaSerGluAsnAsnValGluGlu 140
DB 393 CAGGAGGAGACATCAGAAATTAATTGAACAGATCCCTGCTTCAGAAAAACAATGGAGAA 452
QY 141 MetValGlnProAlaGluSerGlnAlaAsnAspValGlyPheLysLysValPheLysPhe 160
DB 453 ATGGTACACCTGTGAGTCCAGGCTTAATGATGTGGCTTCAGAAAGATTAATTT 512
QY 161 ValGlyPheLysPheThrValLysLysAspLysAsnGluLysSerAspThrValGlnLeu 180
DB 513 GTTGGTTTAATTCACGATGAAGAAAGATTAATAAAGTCAATCTGTCCAACTA 572
QY 181 LeuThrValLysLysAspGluGlyGluGlyAlaGluAlaSerValGlyAlaGlyAspHis 200
DB 573 CTCACGTGCAAGAGATAGAGCGAAGGGGCGAGAACCTCTGTCCGAGCTGGAGACCA 632
QY 201 GlnGluProSerValGluThrAlaValGlyGluSerAlaSerLysGluSerGluLeuLys 220
DB 633 CAGGAGCCGAGTGTGAGACTGCGGTGAGAGATCAGATCCAAAGAAAGTGAAG 692
QY 221 GlnSerThrGluLysGlnGluGlyThrLeuLysGlnGluGlnSerSerThrGluLeuPro 240
DB 693 CAATCCACAGAGAGCAAGAACCCCTGAAGCAAGAACAGAGCAAGCAAGAAATCCCC 752
QY 241 LeuGlnAlaGluSerAspGlnAlaAlaGluGluGluAlaLysAspGluGlyGluGluLys 260
DB 753 CTTCAAGCCGAATCTGATCAAGCGCTGAGGAAGAACCAAGATCAAGAGAGAAAAA 812
QY 261 GlnGluLysGluProThrLysSerProGluSerProSerSerProAlaAsnSerGluThr 280
DB 813 CAGAGAAAGAGCCCAAGTCCCAAGATCCCGAGAGGCCAGCTCAACAGTGAAGCA 872
QY 281 ThrSerSerPheLysLysPhePheThrHisGlyTrpAlaGlyTrpArgLysLysThrSer 300
DB 873 ACATCTTCTTCAAGAAAGTCTTCACTCAGCGTTGGCCGCGTGGGCAAGAAAGACAG 932
QY 301 PheLysLysSerLysGluAspAspLeuGluThrAlaGluLysArgLysGluGlnGluAla 320
DB 933 TTCAGAAATCAAAAAGAGATGATCTGAAACTGCCAGAGAAAGAAAGGAGCAAGGCA 992
QY 321 GlnLysValAspGluGlnGluLysGluLysThrGluProAlaSerGluGluGlnGluPro 340

DB 993 GAAAAAGTAGACGAGAAAGAAAAAGACAGAGCCAGCTCGGAGAGACAGAGCCG 1052
QY 341 AlaGluAspThrAspGlnAlaArgLeuSerAlaAspTrpGluLysValGluLeuProLeu 360
DB 1053 GCAGAAACACAGACACAGGCCAGGTTGACAGACACTACAGAAAGGTGAGACTCTTGG 1112
QY 361 GlnAspGlnValGlyAspLeuGluAlaSerSerGluGluLysCysAlaProLeuAlaThr 380
DB 1113 GAAAGCACGTTGGTGTACCTGAGGCATGTCAAGAGAAAGTGTGCTCTTTGGCAACG 1172
QY 381 GluValPheAspGluLysMetGluAlaHisGlnGluValValAlaGluValHisValSer 400
DB 1173 GAAGTGTGATGAGAAAGATGAAGCCACCAAGAGATGTGTGACAGGCTCCAGTGAGC 1232
QY 401 ThrValGluLysThrGluGluGlnGlyGlyGlyGluAlaGluGlyGlyValAla 420
DB 1233 ACCGTGAGAAAGACAGAGAGAGACAGAGAGAGAGAGAGGCTGAAGGGGCGTGTG 1292
QY 421 ValGluGlyThrGlyGluSerLeuProGluLysLeuAlaGluProGlnGluValPro 440
DB 1293 GTAGAAAGAAAGAGAAATCTTGCCCTTGAGAACTGGCTGAGCCCGAGAGGTCCTCC 1352
QY 441 GlnGluAlaGluProAlaGluGluLeuMetLysSerArgGluMetCysValSerGlyGly 460
DB 1353 CAGGAAGCTGAGCTGTGAGAGAGCTGATGAAGAGCAGAGAGATGTGTCTTGAGAGA 1412
QY 461 AspHisThrGlnLeuThrAspLeuSerProGluGluLysThrLeuProLysHisProGlu 480
DB 1413 GACCACTCAACTGACAGACTTAAGTCTGAAGAGAAAGAGCTGCCCAACACCAADA 1472
QY 481 GlyLeValSerGluValGluMetLeuSerSerGlnGluArgL1eLysValGlnGlySer 500
DB 1473 GCATTTGCAAGTGAAGTGAAGTGTCTCTTCAGAAAGAAATCAAGTCAAGGAAGT 1532
QY 501 ProLeuLysLysLeuPheSerSerSerGlyLeuLysLysLeuSerGlyLysLysGlnLys 520
DB 1533 CCTTGAAAGAACTCTTCAAGTATGCTCAGGCTTAAAGAGCTGTCTGGAGAAAGCAGAAG 1592
QY 521 GlyLysArgGlyGlyGlyLysAspGluGluProGlyGlyLysGlnHisL1eHisThrGlu 540
DB 1593 GGGAAACAGAGGAGGTGGGGAGACAGAAAGCTGGAGAAATCAACATTCACACCCAA 1652
QY 541 SerProGluSerAlaAspGluGlnLysGlyGluSerSerAlaSerSerProGluGluPro 560
DB 1653 TCCCAAGAGAGTGTGAGAGAGAGAAAGGAGAGCTGTGCTGCCCGAGAGAGCTT 1712
QY 561 GlnGluThrThrCysLeuGluLysGlyProLeuGluAlaProGlnAspGlyGluAlaGlu 580
DB 1713 GAGAGACCACTGTGTGAGAAAGGCCCTGGAAGACACCCCGAGATGGGGAAGCTAG 1772
QY 581 GlnGlyThrThrSerAspGlyGluLysLysArgGluLysL1eThrProTrpAlaSerPhe 600
DB 1773 GAAGAACTACTTCCGATGGAGAGAAAGAGAGAGAGATCACTCCCTGGGCACTCTTC 1832
QY 601 LysLysMetValThrProLysLysArgValArgArgProSerGluSerAspLysGlnGlu 620
DB 1833 AAAAAGATGTGACACCCAGAGAAACGGGTCCGAAACCTTCTGAGAGTGAACAAGAGAA 1892
QY 621 GlnLeuGluLysValLysSerAlaThrLeuSerSerThrAspSerThrValSerGluMet 640
DB 1893 GAGCTGAGAAAGTCAAGAGGCCACCTTGTCTCCACTGATAGCACAGTGTCAAAATG 1952
QY 641 GlnAspGluValLysThrValGlyGluGlnGluLysProGluGluProLysArgArgVal 660
DB 1953 CAAGATGAAGTCAAAACTGTGTGAGGAAACAAAAGCCAGAGAAACCAAAGGTAGG 2012
QY 661 AspThrSerValSerTrpGluAlaLeuL1eCysValGlySerSerLysLysArgArgAla 680
DB 2013 GATACTTCAGTGTCTTGGAGACACTGATTTGTGTGGAATATCCAAAGAGAGCAAGG 2072
QY 681 LysAlaSerSerSerAspAspGluGlyGlyProArgThrLeuGlyGlyAspSerHisArg 700
DB 2073 AAGGCATCTCTTCAAGATGATGAAGAGGGCCAAAGACACTGGAGGGGACAGTACAGA 2132

QY 701 AlaGluAlaSerLysAspLysGluAlaGlyThrAspAlaValProAlaSerThrGln 720
DB 2133 GCAGAGAGCCGCAAGAACAAAGAGAGCCGGAACAGACGCTGTCTCTGCAGACCCAG 2192
QY 721 GluGlnAspGlnAlaGlnGlySerSerProGluProAlaGlySerProSerGluGly 740
DB 2193 GAGCAGGACCAAGCGCAAGAGGTTCTCCACCGAGCCAGCGGAGCCCTTCCGAAGGG 2252
QY 741 GluGlyValSerThrTrpGluSerPheLysArgLysValThrProArgLysLysSerLys 760
DB 2253 GAAGGTGTCTCCACTTGGGAGTCATTTAAAGATTAGTCACTCCAAGAAAAAATCCAAG 2312
QY 761 SerLysLeuGluGluLysAlaGluAspSerSerValGluGlnLeuSerThrGluIleGlu 780
DB 2313 TCAAACTGGAAGAGAAAGCCGAGACTCTAGTGTAGAGCAGTTGTCCACTGAGATCGAA 2372
QY 781 ProSerArgGluGluSerTrpValSerIleLysLysPheIleProGlyArgArgLysLys 800
DB 2373 CCAGTAGAGAGAAATCTTGGGTTTCCATTAGAAATTCTATCCCGGACCGCGGAAGAAA 2432
QY 801 ArgAlaAspGlyLysGlnGluGlnAlaThrValGluAspSerGlyProValGluIleAsn 820
DB 2433 AGGCGAGCGGAGCAAGAACAGCCACTGTGGAAGACTCAGGCGCCAGTGGAGATAAT 2492
QY 821 GluAspAspProAsnValProAlaValValProLeuSerGluTyrAsnAlaValGluArg 840
DB 2493 GAGGACGACCCCTAATGTCCAGCGCTGTGCTCTGAGTATATGCGAGTGCAGAGG 2552
QY 841 GluLysMetGluAlaGlnGlyAsnThrGluLeuProGlnLeuGlyAlaValTyrVal 860
DB 2553 GAGAAGATGAAGCCCAAGGGAATACGGAGCTGCCCCAGCTGTGGGGGCTGTGTACGTG 2612
QY 861 SerGluGluLeuSerLysThrLeuValHisThrValSerValAlaValIleAspGlyThr 880
DB 2613 TCCGAGGAGCTCAGTAGACTCTGGTCCACACTGTGAGTGTGCGAGTCATTGATGGACC 2672
QY 881 ArgAlaValThrSerValGluGluArgSerProSerTrpLysSerAlaSerValThrGlu 900
DB 2673 AGGCGAGTCAACAGTGTGCAAGAGCGGTCTCCTTGTGGATATCCGCTTCCGTAACAGAA 2732
QY 901 ProLeuGluHisThrAlaGlyGluAlaMetProProValGluGluValThrGluLysAsp 920
DB 2733 CTTCTTGAACACACAGCGGAGAGCCATGCCACCTGTTGAAGAGGTCACTGAAAAAGAC 2792
QY 921 IleIleAlaGluGluThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHis 940
DB 2793 ATCATTCAGAGAAACTCTCTGTCTCACCAGACGTTACAGAGGGTAAAGATGCCCAT 2852
QY 941 AspAspMetValThrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThr 960
DB 2853 GACGACATGGTCAACAGTGAAGTGAATTTCACTCTCAGAAAGCTGTGACAGCCACAGAGACC 2912
QY 961 SerGluAlaLeuArgThrGluGluValThrGluAlaSerGlyAlaGluGluThrThrAsp 980
DB 2913 TCAGAGGCTCTCCGTAAGAGTTTACCAGAAAGTTCGGGGGCCGGAAGACCCACAGAC 2972
QY 981 MetValSerAlaValSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrPro 1000
DB 2973 ATGGTGTCCGAGTTTCCAGCTGACTGACTTCCCGACACACCACAGAGGAAGCCACCCCA 3032
QY 1001 ValGlnGluValGluSerGlyValLeuAspThrGluGluGluGluArgGlnThrGlnAla 1020
DB 3033 GTTCAGGAGGTAGAGAGTGGTGTCTAGATACAGAAAGAGAGGCGCCAGAGCGAGGCC 3092
QY 1021 IleLeuGlnAlaValAlaAspLysValLysGluGluSerGlnValProAlaThrGlnThr 1040
DB 3093 ATCCTCCAAGCCGTTCAGACAAAGGTGAAAGAGGAGTCCCGAGGTGCTCGCAACCCAGACT 3152
QY 1041 ValGlnArgThrGlySerLysAlaLeuGluLysValGluGluValGluGluAspSerGlu 1060
DB 3153 GTCCAGAGAAACGGGTCAAAGCACTGGAGAGGTTGAGGAGGTAGAGGAGCTCCGAA 3212

QY 1061 ValLeuAlaSerGluLysGluLysAspValMetProLysGlyProValGlnGluAlaGly 1080
DB 3213 GTGCTGGCTTCGGAGAAAGAGAGGAGCTTATGCCGAAAGAGCCCGTGCAGGAAGCTGGA 3272
QY 1081 AlaGluHisLeuAlaGlnGlySerGluThrGlyGlnAlaThrProGluSerLeuGluVal 1100
DB 3273 GCTGAGCATCTTGCACAGGGCTCTGAGACTGGACAGGCTACTCCAGAGAGCCTTGAAGTT 3332
QY 1101 ProGluValThrAlaAspValAspHisValAlaThrCysGlnValIleLysLeuGlnGln 1120
DB 3333 CCTGAAGTCACGGCAGATGTAGACCATGTGCCACGTCGCGAGTTATCAAGCTCCAGCAG 3392
QY 1121 LeuMetGluGlnAlaValAlaProGluSerSerGluThrLeuThrAspSerGluThrAsn 1140
DB 3393 CTGATCGAACAGGCGCTGGCCCTGAGTCTATCCGAAACCTTGACACAGTGAAGCAAT 3452
QY 1141 GlySerThrProLeuAlaAspSerAspThrAlaAspGlyThrGlnGlnAspGluThrIle 1160
DB 3453 GGNAGCACTCCCTTAGCAGATTGACACACTGCAGATGGGACACAGCAAGATGAACCATT 3512
QY 1161 AspSerGlnAspSerLysAlaThrAlaValArgGlnSerGlnValThrGluGluGlu 1180
DB 3513 GACAGCCAGCAGATAAAGCCACTGCAGCTGTCCAGGCTCAGGCTCAGCAAGAAGAG 3572
QY 1181 AlaAlaThrAlaGlnLysGluGluProSerThrLeuProAsnAsnValProAlaGlnGlu 1200
DB 3573 GCGGCTACTGCTCAGAAAGAGGAGCCTTCGACACTACCTAATAATTTCCAGCCCCAGGAA 3632
QY 1201 GluHisGlyGluGluProGlyArgAspValLeuGluProThrGlnGlnGluLeuThrAla 1220
DB 3633 GAACATGGGGAGAAACAGAAAGAGATGTTCTTGAACCTTACACAGCAAGAGCTTACTGCT 3692
QY 1221 AlaAlaValProValLeuAlaLysThrGluValGlyGlnGluGlyGluValAspTrpLeu 1240
DB 3693 GCAGCGCTGCCGCTTCTGCCAAAGACTGAGGTGGGTCAAGAGGGTCAGGTTGACTGGTTG 3752
QY 1241 AspGlyGluLysValLysGluGluGlnGluValPheValHisSerGlyProAsnSerGln 1260
DB 3753 GATGGAGAAAAAGTCAAAGAAAGACAGGAGGTGTTTTGTACACTCTGGACCCCAACAGTCAA 3812
QY 1261 LysAlaAlaAspValThrTyrAspSerGluValMetGlyValAlaGlyCysGlnGluLys 1280
DB 3813 AAGCTCTGTGTGTGACATATGACATGAGTGTGAGTGGAGTGGCCGGGTGTCAAGGAAAG 3872
QY 1281 GluSerThrGluValGlnSerLeuSerLeuGluGluGlyGluMetGluThrAspValGlu 1300
DB 3873 GAGAGTACTGAAGTCAGAGTCTTACCTGGAGGAGGAGAGATGAAACTGACGTTGAA 3932
QY 1301 LysGluLysArgGluThrLysProGluGlnValSerGluGluGlyGluGlnGluThrAla 1320
DB 3933 AAGGAGAAAGGGAGACAAAGCCAGAGCAAGTGAGTGAAGAGGTGAGCAGGAAACAGCC 3992
QY 1321 AlaProGluHisGluGlyThrTyrGlyLysProValLeuThrLeuAspMetProSerSer 1340
DB 3993 GCTCCTGAGATGAAGAAACCTACGGGAGCCAGTCTGACACTTGACATGCCAGCTCA 4052
QY 1341 GluArgGlyLysAlaLeuGlySerLeuGlyLysProSerLeuProAspGlnAspLys 1360
DB 4053 GAGAGGGGAGAGCACTGGGAGGCTTGGAGGAGGAGCCCTTCTCCAGACCAAGACAAA 4112
QY 1361 AlaGlyCysIleGluValGlnValGlnSerLeuAspThrValThrGlnThrAlaGlu 1380
DB 4113 GCAGGTTCATAGAGGTTCAAGTTCAAGCCTGGAACAACAGTCACTCAAAACAGCAGAA 4172
QY 1381 AlaValGluLysValIleGluThrValValIleSerGluThrGlyLysProGluCys 1400
DB 4173 GCTGTGAAAAGGTCTATAGAAACGGTTGTGATTTACAGACAGGTGAAAGTCCAGAGTGT 4232
QY 1401 ValGlyAlaHisLeuLeuProAlaGluLysSerSerAlaThrGlyGlyHisTrpThrLeu 1420
DB 4233 GTAGGTGCACACTTATACAGCTGAGAGTCTCTCTGCAACGGGTGGCCACTGACTCTT 4292
QY 1421 GlnHisAlaGluAspThrValProLeuGlyProGluSerGlnAlaGluSerIleProfile 1440

QY 708 sGluAlaGlyThrAspAlaValProAlaSerThrGlnGluGlnAspGlnAlaGlnGlySe 728
DB 1133 AGAGCCGA-ACAGACGCTGTTCTGCGACACCCAGAGAGGACCAAGCGCAAGGAAG 1191
QY 728 rSerSerProGluProAlaGlySerProSerGluGlyGluGlyValSerThrTrpGluSe 748
DB 1192 TTCTCACCAGGAGCAGCGGAGCCCTTCGAAGGGAGGAGGTCTCCACTTGGGAGTC 1251
QY 748 rPheLysArgLeuValThrProArgLysLysSerLysSerLysLeuGluLysAla-- 767
DB 1252 ATTTAAAGATTAGTCACTCCAGAAAATAATCCAAGTCAAAACTGGAAGAGAAGA-AG 1310
QY 768 ----GluAspSerSerVal---GluGlnLeuSerThrGluLeuGluProSerArgGluG1 785
DB 1311 CCGAAGGACTCTAGTTGTAGGACGAGCTGTCACCTGAGATCGAACCGGTGAGAGA 1370
QY 785 uSerTrpValSerIleLysLysPheIleProGlyArgArgLysLysArgAlaAspGlyLY 805
DB 1371 ATCTTGGGTTTCCATTAAAGAAATTCATCCCGGACGCGGAGAAAGAGGAGATGGAA 1430
QY 805 s-GlnGluGlnAlaThrValGluAspSerGlyProValGluIleAsnGluAspAspProA 825
DB 1431 GGCAAGAACCAAGCCACTGTGTGAAGACTCAGGGCCAGTGGAGATAAATGAGACGAGCCTG 1490
QY 825 snValProAlaValValProLeuSerGluTyrAsnAlaValGluArgGluLysMetGluA 845
DB 1491 ATGTCCAGCGCTGTCCTGCTGAGTATGATGAGTGGAGGGAGAGATGGAA 1550
QY 845 laGlnGlyAsnThrGluLeuProGlnLeuGlyAlaValTyrValSerGluGluLeuS 865
DB 1551 CCCAGGGGAATGCGGAGCTGCCAGCTGCTGGGCTGTGTA--GTGTCCGAGGAGCTCA 1607
QY 865 erLysThrLeuValHisThrValSerValAlaValIleAspGlyThrArgAlaValThrS 885
DB 1608 GTAAGACTCTGGTCCACACTGTGAGTGTGCGAGTCAATTGATGGACCGAGGCGAGTCA 1667
QY 885 erValGluGluArgSerProSerTrpIleSerAlaSerValThrGluProLeuGluHisT 905
DB 1668 GTGTGGAAGAGCGGTCTCTTGTGGATATCCGCTTCGTAACAGAACCTCTTGAACACA 1727
QY 905 hrAlaGlyGluAlaMetProProValGluGluValThrGluLysAspIleAlaGluG 925
DB 1728 CAGCGGAGAAGCCATGCCACCTGTTGAAGAGGTCACTGAAAGAGACATATTGCGAGAAG 1787
QY 925 luThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHisAspAspMetValT 945
DB 1788 AAACCTCTGTGCTCACCCAGACGTTTACAGAGGGTAAAGATGCCATGACGACATGTGTC 1847
QY 945 hrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeuA 965
DB 1848 CCAGTGAAGTGGATTTTCACTCAGAAGCTGTGACAGCCACAGAGACCTCAGAGGCTCTCC 1907
QY 965 rgThrGluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSerAlav 985
DB 1908 GTACTGAAGAAGTTTACCAGAGCATCGGGGCGGAGAGACCAAGACATGGTGTCCGAG 1967
QY 985 alSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGluValG 1005
DB 1968 TTTCCAGAGTGAAGTCTCCAGACACCAAGAGGAGCCACCCAGTTTCCAGAGGTAG 2027
QY 1005 luSerGlyValLeuAspThrGluGluGluArgGlnThrGlnAlaIleLeuGlnAlav 1025
DB 2028 AGGTGTGTGTGTAGTACAGAGAAGAGGAGGCGCCAGAGCGAGGCTCCTCCAGAGCCG 2087
QY 1025 alAlaAspLysValLysGluGluSerGlnValProAlaThrGlnThrValGlnArgThrG 1045
DB 2088 TTGCAGACAGGTGAAGAGAGATCCCGAGTGTGCTGCAACCCAGACTGTGCAGAGAAGCG 2147
QY 1045 lySerLysAlaLeuGluLysValGluGluValGluGluAspSerGluValLeuAlaSerG 1065
DB 2148 GGTCAAAGCACTGGAGAGGTTGAGGAGGTAGAGGAGGACTCCGAAGTGTCTGGCTTCGG 2207
QY 1065 luLysGluLysAspValMetProLysGlyProValGlnGluAlaGlyAlaGluHisLeuA 1085

DB 2208 AGAAAGAGAGGACGTTATGCCCCAAAGGACCCCTGCAGGAAGCTGGAGCTGACATCTTG 2267
QY 1085 laGlnGlySerGluThrGlyGlnAlaThrProGluSerLeuGluValProGluValThrA 1105
DB 2268 CACAGGGCTCTGAGACTGGACAGGCTACTCCAGAGAGCCCTTGAAGTTCTTGAAGTCACAG 2327
QY 1105 laAspValAspHisValAlaThrCysGlnValIleLysLeuGlnLeuMetGluGlnA 1125
DB 2328 CAGATGTAGACCATGTGCGCACGTGCCAGGTATCAAGCTCCAGCAGCTGATGAAACAGG 2387
QY 1125 laValAlaProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrProL 1145
DB 2388 CCGTGCCCTCTGAGTCAATCCGAAACCTTGACACAGACTGAGACAAATGGAAGCAGCTCCCT 2447
QY 1145 euAlaAspSerAspThrAlaAspGlyThrGlnGlnAspGluThrIleAspSerGlnAspS 1165
DB 2448 TAGCAGATTACAGACACTGCAGATGGGACACAGCAAGATGAACCATTTGACAGCCAGGACA 2507
QY 1165 erLysAlaThrAlaAlaValArgGlnSerGlnValThrGluGluGluAlaAlaThrAlaG 1185
DB 2508 GTAAGCCCACTGCAGCTGTGAGGAGTCAAGGTCACAGAGTCAAGAGAGGCGGCTACTGCTC 2567
QY 1185 lnLysGluGluProSerThrLeuProAsnValProAlaGlnGluGluHisGlyGluG 1205
DB 2568 AGAAAGAGGAGCCCTTCGACACTACCTAAATGTTCCAGCCAGGAAGAACATGGGGAAG 2627
QY 1205 luProGlyArgAspValLeuGluProThrGlnGlnGluLeuThrAlaAlaAlaValProV 1225
DB 2628 AACCCGGAAGAGATGTTCTTGAACCTACACAGCAAGAGCTTGTCTGCAGCCGCTGCCCG 2687
QY 1225 alLeuAlaLysThrGluValGlyGlnGluGlyGluValAspTrpLeuAspGlyGluLysV 1245
DB 2688 TCTGGCAAAAGACTGAGGTGGGTCAAGAGGTGAGGTGACTGTTGGATGGAGAAAAG 2747
QY 1245 alLysGluGlnGluValPheValHisSerClyProAsnSerGlnLysAlaAlaAspV 1265
DB 2748 TCAAGAAAGAACAGGAGGTGTTGTACACTCTGGACCCCAACAGTCAAAAGGCTGCTGATG 2807
QY 1265 alThrTyrAspSerGluValMetGlyValAlaGlyCysGlnGluLysGluSerThrGluV 1285
DB 2808 TGACATATCACAGTGAAGTGATGGAGTGGCGGTGTCCAGGAAAGAGAGTACTGNAAG 2867
QY 1285 alGlnSerLeuSerLeuGluGluGlyGluMetGluThrAspValGluLysGluLysArgG 1305
DB 2868 TGACAGAGTCTTAGCTTGGAGGAGGAGAGTGAACCTGACGTTGAAAGAGGAGAAAAGGG 2927
QY 1305 luThrLysProGluGlnValSerGluGluGlyCyluGlnGluThrAlaAlaProGluHisG 1325
DB 2928 AGCAAAAGCCAGAGCAAGTGAGTGAAGAGGTGAGCAGGAACACAGCCGCTCTCTGAGCATG 2987
QY 1325 luGlyThrTyrGlyLysProValLeuThrLeuAspMetProSerSerGluAtgGlyLysA 1345
DB 2988 AAAGGAACCTACGGGAAGCCAGTCTCTGACACTTGACATGCCAGCTCAGAGAGGGGAGAG 3047
QY 1345 laLeuGlySerLeuGlyGlySerProSerLeuProAspGlnAspLysAlaGlyCysIleG 1365
DB 3048 CACTGGGAAGCCTTGGAGGAAGCCCTTCTCTCCAGACCAAGACAAAGCAGGTTGTCATAG 3107
QY 1365 luValGlnValGlnSerLeuAspThrThrValThrGlnThrAlaGluAlaValGluLysV 1385
DB 3108 AGGTTCARGTTCAAAGCCCTGGACACACAGTCACTCAAAACAGCAGGAAGCTGTGGAAAGG 3167
QY 1385 alIleGluThrValValIleSerGluThrGlyLysSerProGluCysValGlyAlaHisL 1405
DB 3168 TCATAGAACAGGTTGTGATTTTCAGACAGAGGTGAAGTCCAGAGTGTGTAGGTGCACACT 3227
QY 1405 euLeuProAlaGluLysSerSerAlaThrGlyHisTrpThrLeuGlnHisAlaGluA 1425
DB 3228 TATTACCACTGAGAAGTCTCTCTGCAACGGGTGGCCACTTGGACTCTTCAGCATGCGAGG 3287
QY 1425 spThrValProLeuGlyProGluSerGlnAlaGluSerIleProIleIleValThrProA 1445

Db 1060 CAAGCCACCACAGGAGCGGCAGAAAGTGCCTCCACAGAGCCCGGTTATCAGCTGAATAT 1119
Qy 354 GluLysValGluLeuProLeuGluAspGlnValGlyAspLeuGluAlaSerSerGluGlu 373
Db 1120 GAGAAAGTTGAGCTGCCCTCAGAGGAGCAAGTCAGTGGCTCGCAGGGACCTTCTGAAGAG 1179
Qy 374 LysCysAlaProLeuAlaThrGluValPheAspGluLysMetGluAlaHisGln---Glu 392
Db 1180 AAACCTGCTCCGTGGCGCAGAGAAGTGTGATGAGAAAATAGAAAGTCCCAAGAGAGAG 1239
Qy 393 ValValAlaGluValHisValSerThrValGluLysThrGluGluGlnGluGlyGly 412
Db 1240 GTTGTGGCCGAGTCCACGTCAGCACCGCTGAGAGAGAACCGAGAGCAG----- 1290
Qy 413 GlyGluAlaGluGlyGlyValValGluValGluGlyThrGlyGluSerLeuProGluLys 432
Db 1291 -----AAACCGAGGTGGAAGAAACAGCAGGGTCTGTGCCAGCTGAAGAA 1335
Qy 433 LeuAlaGluProGlnGluValProGlnGluAlaGluProAlaGluGluLeuMetLysSer 452
Db 1336 TTGGTTGGAATGGATGCAGAACCTCAGAAAGCCGAACCTCCCAAGGAGCTGGTGAAGCTC 1395
Qy 453 ArgGluMetCysValSerGlyGlyAspHisThrGlnLeuThrAspLeuSerProGluGlu 472
Db 1396 AAAGAAACGTGTGTTCCGAGAGGACCCTACAGAGGAGCTGACCTCAGTCTCTGATGAG 1455
Qy 473 LysThrLeuProLysHisProGluGlyValSerGluValGluMetLeuSerSerGln 492
Db 1456 AAGTGTCTGCCAAACCCCGAAGGGGTTGTGAGTGAGGTGGAATGCTGCATCACAG 1515
Qy 493 GluArgIleLysValGlnGlySerProLeuLysLysLeuPheSerSerGlyLeuLys 512
Db 1516 GAGAGATGAAGTGCAGGGAAAGTCCACTAAAGAAGCTTTTACCAGCACTGGCGTTAAAA 1575
Qy 513 LysLeuSerGlyLysLysGlnLysGlyLysArgGlyGlyGlyGlyAspGluGluProGly 532
Db 1576 AAGCTTTCTGGAAGAAACAGAAAGGGAAGA---GGAGGAGAGCAGAGGAATCAGGG 1632
Qy 533 GlyTyrGlnHisIleHisThrGluSerProGluSerAlaAspGluGlnLysGlyGluSer 552
Db 1633 GAGCACACTCAGGTTCCAGCGGATTCTCCGACAGCCAGGAGGAGCAAAAGGGCGAGAGC 1692
Qy 553 SerAlaSerSerProGluGluProGluGluThrThrCysLeuGluLysGlyProLeuGlu 572
Db 1693 TCTGCTCATCCCTCAGGAGCCGAGGAGATCACGTGCTCTGGAAGAGGCGTTAGCCGAG 1752
Qy 573 AlaProGlnAspGlyGluAlaGluGluGlyThrThrSerAspGlyGluLysLysArgGlu 592
Db 1753 GTCAGCAGGATGGGGAAGCTGAAGAAGGAGCTACTTCCGATGGAGAGAAAAGAGAA 1812
Qy 593 GlyIleThrProTrpAlaSerPheLysMetValThrProLysLysArgValArg 612
Db 1813 GGTGTCACTCCCTGGGATCATTTCAAAAGAGTGGTGAGCGCCCAAGAGCGTGTAGACGG 1872
Qy 613 ProSerGluSerAspLysGluGluLeuGluLysValLysSerAlaThrLeuSerSer 632
Db 1873 CCTTCGAAAGTGATAAAGAGATGAGCTGCACAAGGTCAAGAGCGCTACTCTGTCTTCC 1932
Qy 633 ThrAspSerThrValSerGluMetGlnAspGluValLysThrValGlyGluGluGlnLys 652
Db 1933 ACCGAGAGCAGCCTCTGAATGCAAGAGAAATGAAGGGAGCGGTGGAAGAGCCCAAG 1992
Qy 653 ProGluGluProLysArgValAspThrSerValSerTrpGluAlaLeuIleCysVal 672
Db 1993 CCGGAGAACCAAGCCAGAGTGGATACCTCAGTATCTTGGAGAGCTTTAATTTGTGTG 2052
Qy 673 GlySerSerLysLysArgAlaArgLysAlaSerSerSerAspGluGlyGlyProArg 692
Db 2053 GGATCATCCAAGAAAAGAGCAAGAGAGGTCTCTTCTCATGAGGAAGGGGAGCCAAAA 2112
Qy 693 ThrLeuGlyGlyAspSerHisArgAlaGluAlaSerLysAspLysGluAlaGlyThr 712
Db 2113 GCAATGGGAGGAGACCACCAAGAAAGCTGATGAGCGCCGGAAGAAAGCAAGAGACGGGACA 2172

Qy 713 AspAlaValProAlaSerThrGlnGluGlnAspGlnAlaGlnGlySerSerSerProGlu 732
Db 2173 GACGGGATCCTTGCTGTTCCCAAGAACATGATCCAGGGCAGGGAAGTCTCTCCCGGAG 2232
Qy 733 ProAlaGlySerProSerGluGlyGlyValSerThrTrpGluSerPheLysArgLeu 752
Db 2233 CAAGCTGGGAAGCCCTACCGAAGGGGAGGCGTTTCCACCTGGGAGTCATTTAAAGGTTA 2292
Qy 753 ValThrProArgLysLysSerLysSerLysLeuGluGluLysAlaGluAsp----- 769
Db 2293 GTCACGCCCAAGAAAAAATCAAGTCCAGCTGGAAGAGAAAAGCGAAGCTCCATAGCT 2352
Qy 770 ---SerSerValGluGlnLeuSerThrGluLeuGluProSerArgGluGluSerTrpVal 788
Db 2353 GGGTCTGGTGTAGAACATTTCCACTCCAGACACTGAAACCCGGTAAAGAAATCCTGGGTC 2412
Qy 789 SerIleLysLysPheIleProGlyArgArgLysLysArgAlaAspGlyLysGlnGluGln 808
Db 2413 TCAATCAAGAGATTTATCTCTGACGAAAGAAAGGCCAGATGGGAAAACAAAGAACAA 2472
Qy 809 AlaThrValGluAspSerGlyProValGluIleAsnGluAspAspProAsnValProAla 828
Db 2473 GCCCTGTTGAAGACGACGAGGCCACAGGGGCCAACGAAAGTGACTCTGTATGTTCCGGCC 2532
Qy 829 ValValProLeuSerGluTyrAsnAlaValGluArgGluLysMetGlu-----Ala 845
Db 2533 GTGGTCCCTCTGTCTCAGTATGATGCTGTAGAAAGGGAGAAAATGAGGCACAGCAAGCC 2592
Qy 846 GlnGlyAsnThrGluLeuProGlnLeuLeuGlyAlaValTyrValSerGluGluLeuSer 865
Db 2593 CAAAAAGGCGCAGACGAGCCGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2652
Qy 866 LysThrLeuValHisThrValSerValAlaValIleAspGlyThrArgAlaValThrSer 885
Db 2653 GAGAGTCAGTTTATGATGAGCAGCAGCTGTCTGCTTCAAGTGCAGAACCTCTTGAACAAGTA 2712
Qy 886 ValGluGluArgSerProSerTrpIleSerAlaSerValThrGluProLeuGluHisThr 905
Db 2713 ATTTGAAGAAAGGTCTCTTCTTGGATATCTGCTTCAAGTGCAGAACCTCTTGAACAAGTA 2772
Qy 906 AlaGlyGluAlaMetProProValGluGluValThrGluLysAspIleAla---Glu 924
Db 2773 GAAGCTGAAGCGCACTGTTAACTGAGGAGGTATTGGAAGAGAGAAATTTGAGAGAA 2832
Qy 925 GluThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHisAspAspMetVal 944
Db 2833 GAACCCCCACGGTTACTGAACCTCTGCAGAGAACAGAGAGCCCGGGCGACAGCGTC 2892
Qy 945 ThrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeu 964
Db 2893 GTTAGTGAGCGGAATTGACCCCGGAGCTGTGACAGCTGCAGAACTGCAGGGCCATTG 2952
Qy 965 ArgThrGluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSerAla 984
Db 2953 GGTTCGAAAGAGAAACCGAAGCATCTGCTGTGGAAGAGACCAAGAAATTTGTTGTCAGCA 3012
Qy 985 ValSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGluVal 1004
Db 3013 GTCCTCCAGTTAAACCGACTCCCGACACACAGAGAGGCGCACTCCGTTGAGGAGGTG 3072
Qy 1005 GluSerGlyValLeuAspThrGluGluGluArgGlnThrGlnAlaIleLeuGlnAla 1024
Db 3073 GAAGTGGCGGTACTGCATAGAGAGCAAGAGAGCGGAGCTCAAGAGGTCTCCAGGCA 3132
Qy 1025 ValAlaAspLysValLysGluGluSerGlnValProAlaThr----- 1038
Db 3133 GTGGCAGAAAAAGTGAAGAGGAATCCAGCTGCCTGGCACCGGTGGGCCAGAAAGTGTG 3192
Qy 1039 ---GlnThrValGlnArgThrGlySerLysAlaLeuGluLysValGluGluValGluGlu 1057
Db 3193 CTTTCAACCTGTGCAGAGACA-----GAGGCAGAAAGACCCAGAGAG 3234

QY 1058 AspSerGluValLeuAlaSerGluValAspValMetProGlyProValGlu 1077
 Db 3235 CAGGCTGAAGCCGTGGGTGAGAAAGAGACGATGTGTGTAAGTAACTGCTCAG 3294
 QY 1078 GluAlaGluValGluHisLeuAlaGlnGlySerGluThrGluGlnAlaThrProGluSer 1097
 Db 3295 GAGGCAAAAACGAGCCTTTTACACAGAGAGAGAGGTGTGGGCGACCAACCCACAGAAAC 3354
 QY 1098 LeuGluValProGluValThrAlaAspValAspHisValAlaThrCys 1113
 Db 3355 TTGTAAGAAAGCTCCTCACTGACAGAGACATAGACTCCAGTGTAACTTAACACTTGT 3414
 QY 1114 GlnValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 1127
 Db 3415 CAAAGCGAAACCTTACGCTGGGTAAATACAGAGATGTGTGTAAGAGAGGCTATCCCC 3474
 QY 1128 ProGluSerSerGluThrLeuThrAspSerGluThrAspGlySerThrProLeuAlaAsp 1147
 Db 3475 CCTGACTCGGTGAAACCCCTACAGACATGATGATGAGACACCCCTGAGCCGAC 3534
 QY 1148 SerAspThrAlaAspGlyThrGlnGlnAspGluThrLeuAspSerGlnAspSerVala 1167
 Db 3535 TTGTAAGCAACCGACGACCAACCCAGAAAGACGATGTGGAAATCCATGAGAGAAATGAG 3594
 QY 1168 ThrAlaAlaValArgGlnSerGlnValThrGluGluGluAlaAlaThrAlaGlnGlySer 1187
 Db 3595 GTGCGATCTGTGATCCAGTCAGGGGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3654
 QY 1188 GluProSerThrLeuProAsnValProAlaGlnGluHisGlyGluGluProGly 1207
 Db 3655 AGGCTCCAGCA---CCTTCAGTTTGTGTTCCAGAAAGAACTTAAGAACATCAAAAG 3711
 QY 1208 ---ArgAspValLeuGluProThrGlnGlnGluLeuThrAlaAlaValProValLeu 1226
 Db 3712 ATGGAAGACACTCTAGACATACAGATAAAGAGGTGTGTAAGTGTATCACTTGTG 3771
 QY 1227 AlaValThrGluValGluGlnGluGluValAspThrLeuAspGlyGluValValVal 1246
 Db 3772 TCMAAGACTGAGGGGACTCAA---GAGGCTGACCACTGATGATGAGAAACCAAA 3825
 QY 1247 ---GluGlnGlnGlnValPheValHisSerGly---ProAsnSer 1259
 Db 3826 GACGTCCATTTTTCAGAGAGCTGAGGGGTCTATACACAGGCACTTAACAGTCACTCG 3885
 QY 1260 GlnValAlaAlaAspValThrTyraSerGluValMetGlyValAlaGlyCysGlnGlu 1279
 Db 3886 GAAAGGTCACTGAAGTGTGCTTAAGGTGAAGGAGACAGAAAGCTGAATGAATAAG 3945
 QY 1280 LysGluSerThrGluValGlnSerLeuSerLeuGluGluGluGlu 1294
 Db 3946 GATGATGCTCTTGAAGTCAAGTCAAGTCACTTCCTCCATCCCCCGTGGAGAGAGAG 4005
 QY 1295 MetGluThrAspValGluValGluValGluValGluValGluValGluValGluVal 1314
 Db 4006 ATGTAGTTCATGATGAG 4065
 QY 1315 Gly---GluGlnGluThrAlaAlaProGluHisGluGluGluGluGluGluGluGlu 1333
 Db 4066 AAGCTTGACACAGAAACAGTGTATCCGATCTGAAGAGGTCTGAAGAGAGAGAGAGAG 4125
 QY 1334 ThrLeuAspMetProSerSerGluValGlyValAlaLeuGlySerLeuGlySerLeu 1352
 Db 4126 ACAGTGAATGTCCCATCATATGAGTGGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4185
 QY 1353 ProSerLeuProAspGlnAspValAlaGlyCysAlaGluValGlnValGlnSerLeuAsp 1372
 Db 4186 CCTCCCTGCTAGAGTCAAG 4245
 QY 1373 ThrThrValThrGlnThrAlaGluAlaValGluValValLeu---GluThrValVal 1390
 Db 4246 GCATCATTCATCTTAACAG 4305
 QY 1391 HisSerGluThrGluGluSerProGluCysValGlyAlaHisLeuLeuProAlaGluVal 1410

Db 4306 ATTTAGAAACAGGTGAACCTTGAGGCTTCAGGTGACATTTAGTTCTTGAAAGAGAA 4365
 QY 1411 SerSerAlaThrGluGlyHisIleThrLeuGlnHisAlaGluAspThrValProLeuGly 1430
 Db 4366 TCCTCTGAAAAAATGAAGACTTTGCGGCTCATCCAGGAGAAATGTGTGTCCACAGAG 4425
 QY 1431 ProGluSerGlnAlaGluSerIleProIleIleValThrProAlaProGluSerThrLeu 1450
 Db 4426 CCGGACTGTCCGCAAAATTCACACCACTGATATCTGTACTACCAACCAAAAGGCTTA 4485
 QY 1451 HisProAspLeuGlnGluGluHisSerAlaSerGlnArgValArgSerGluGluGluAsp 1470
 Db 4486 AGTTCGACCTGGAAG 4545
 QY 1471 LysProAspAlaGluProAspAlaAspGlyLysGluSerThrAlaIleGluValLeu 1490
 Db 4546 GAGCAGGTTCCTGTCAGAGAGTCC---AAAGTGAATGTAGCAATTTGAGATTTA--- 4596
 QY 1491 LysAlaGluProGlu---IleLeuGluLeuGluSerLysSerAsnLysIleValLeu 1508
 Db 4597 ---GAGCTGAAATATGGATTTTGGACCTTGAGACCAAAAGAGATMAACTGTCCAA 4650
 QY 1509 AsnValIleGlnThrAlaValAlaAspGlnPheAlaArgThr---GluThrAlaProGluThr 1527
 Db 4651 AACATATTCACAGACCGCGTGTGACAGTTGTGTACTACAGAAAGAGAGAGAGAGAGAG 4710
 QY 1528 HisAlaTyraSerGlnThrGlnValProAlaCysArgLeuAspSerArgGluProAsn 1547
 Db 4711 TTGACGTCTGATGTACAGACCAAGCTCAGCGTATMAAGCTGACAGAGAGAGAGAGAG 4770
 QY 1548 ArgCysTrpThrLys---MetLysAspAla-LysMetLysHisProVal 1562
 Db 4771 CAGGAAACGAGAAAG 4830
 QY 1562 LProGlnProArgGlu---AspLeuGlnValLeuThrValLe 1575
 Db 4831 ACTTCAGCCAAAG 4890
 QY 1575 uGluAlaTrpAlaGlnProArgLysCysLeuProArgLeuGlnLeuLysAlaProVal 1594
 Db 4891 AAAGACATGATGAG 4948

 RESULT 4
 US-09-738-877-1
 ; Sequence 1, Application US/09738877
 ; Patent No.: US20020015970A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Murray, Richard
 ; APPLICANT: Watson, Susan
 ; APPLICANT: Weiss, Stephen J.
 ; APPLICANT: Glynn, Richard
 ; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF ANGIOGENESIS, COMPOSITIONS, AND MET
 ; FILE REFERENCE: A-69806/DJB/JTD
 ; CURRENT APPLICATION NUMBER: US/09/738, 877
 ; PRIOR FILING DATE: 2000-12-15
 ; PRIOR APPLICATION NUMBER: US 09/637, 977
 ; PRIOR FILING DATE: 2000-08-11
 ; PRIOR APPLICATION NUMBER: US 60/148, 425
 ; PRIOR FILING DATE: 1999-08-11
 ; PRIOR APPLICATION NUMBER: PCT/US 00/22061
 ; PRIOR FILING DATE: 2000-08-11
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 1
 ; LENGTH: 6608
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; PUBLICATION INFORMATION:
 ; AUTHORS: Nauert et al.
 ; TITLE: Gravin, an autoantigen recognized by serum from myasthenia gravis
 ; TITLE: patients, is a kinase scaffold protein

JOURNAL: Curr. Biol.
VOLUME: 7
ISSUE: 1
PAGES: 52-62
DATE: 1997-01-01
DATABASE ACCESSION NUMBER: U81607
DATABASE ENTRY DATE: 1997-06-26
US-09-738-877-1

Alignment Scores:

Pred. No.: 1,566-292 Length: 6608
Score: 4300.50 Matches: 967
Percent Similarity: 68.81% Conservative: 189
Best Local Similarity: 57.56% Mismatches: 408
Query Match: 53.27% Indels: 116
DB: 10 Gaps: 36

US-09-902-432-4 (1-1596) x US-09-738-877-1 (1-6608)

QY 1 MetGlyAlaGlySerThrGluGlnArgSerProGluGln---ProAlaGlySerAsp 19
DB 192 ATGGGCGCCGGAGCTCCACCGAGCAGCGCGCGGAGCAGCGCGCGGAGGAGCTCC 251
QY 20 ThrProSerGluLeuValLeuSerGlyHisGlyProAlaAlaGluAlaSer---GlyAla 38
DB 252 ACGCCGCTGAGCCGAGCCGAGCGCGCGGCGCCCTCGCGCGGAGCGCGCGCAGACACC 311
QY 39 AlaGlyAspProAlaAspAla-----AspProAlaThrLysLeuProGlnLysAsnGly 56
DB 312 ACCCGGACCCCGCATCGCTCGGACCCCGCCACCAAGCTCTTACAGAGAAATGGT 371
QY 57 GlnLeuSerSerValAsnGlyValAlaGluGlnGlyAspValHisValGlnGluGluAsn 76
DB 372 CAGCTGTCCACCAATCAATGCGTAGCTGAGCAAGATGAGCTCAGCCTCCAGGAGGGTGAC 431
QY 77 GlnGluGlyGln-----GluGlu 82
DB 432 CTAATGCGCAGAAAGAGAGCCCTGAACGGTCAAGAGCCCTTAACACGCCAGGAGGAAGAA 491
QY 83 GluValValAspGluAspValGlyGlnArgGluSerGluAspValArgGluLysAspArg 102
DB 492 GAAGTCATTGTCAGGAGGTGGACAGAGAGACTCTGAAGATGTGACGGAAGAGACTCC 551
QY 103 ValGluGluMetAlaAlaAsnSerThrAlaValGluAspThrLysAspGlyGlnGlu 122
DB 552 GATAAAGAGATGGCTACTAAGTCAGCGGTGTTTCCAGCATCATCAGATGATGGCGAG 611
QY 123 GluThrSerGluLeuLeuGluGlnGluProAlaSerGluAsnValGluGluMetVal 142
DB 612 GAGAACCGAAAT---ATCGAACAGATTCTTCTCAGAAAGCAATTTAGAGAGCTTAACA 668
QY 143 GlnProAlaGluSerGlnAlaAsnAspValGlyPheLysLysValPheLysPheValGly 162
DB 669 CAACCCACTGAGTCCAGGCTAATGATATTGGATTAAAGAGGTGTTTAATGTTGTTGGC 728
QY 163 PheLysPheThrValLysLysAspLysAsnGluLysSerAspThrValGlnLeuLeuThr 182
DB 729 TTTAAATTAAGTGAAGAAAGGATAAGACAGAGAGAGCTGACACTGTCTCCAGCTACTCACT 788
QY 183 ValLysLysAspGluGlyGluGlyAlaGluAlaSerValGlyAlaGlyAspHisGlnGlu 202
DB 789 GTGAAGAAAGATGAAGGGGAGGAGCA-----GCAGGGGCTGGCGACCCACGAGAC 839
QY 203 ProSerValGluThrAlaValGlyGluSerAlaSerLysGluSerGluLeuLysGlnSer 222
DB 840 CCCAGCCTT-----GGGGCTGGAGAGACGATCCCAAGAAAGCAACCAACAATCT 893
QY 223 ThrGluLysGlnGluGlyThrLeuLysGlnGluGlnSerSerThrGluLeuProLeuGln 242
DB 894 ACAGAGAAACCCGAGAGACCTGTAAGCGGTGAGCAAGCCACGACGAGAAATTTCTCCCCCA 953
QY 243 AlaGluSerAspGlnAlaAlaGluGluAlaLysAspGluGlyGluGluLysGlnGlu 262

DB 954 GCCGAATCTGGCCCAACGACGTG---GAGGAATGCAAAAGAGAGGAGAAACAGAA 1010
QY 263 LysGluProThrLysSerProGluSerProSerSerProValAsnSerGluThrThrSer 282
DB 1011 AAAGAACCTAGCAAGTCTGCAGAAATCTCCGAGTAGTCCCGTAGCAGTGAACAGGATCA 1070
QY 283 SerPheLysLysPhePheThrHisGlyTTPAlaGlyTTPArgLysLysThrSerPheLys 302
DB 1071 ACCTTCAAAAAATTTCTCACTCAAGGTTGGCGCGGTGGCGCAAAAGACCACTTCAGG 1130
QY 303 LysSerLysGluAspAspLeuGluThrAlaGluLysArgLysGluGlnGluAlaGluLys 322
DB 1131 AACCCAGAGAGATGAAGTGGAGCTTTCAGAGAGAAAGAAAGAACAGACCCAGAGAAA 1190
QY 323 ValAspGluGluLysGluLysThrGluProAlaSerGluGlu----- 337
DB 1191 GTAGACACAGAAAGACGGAAGGCGAGAGTTGCTCCGAGAAACTGACCGCTCCGAG 1250
QY 338 -----GlnGluProAlaGluAspThrAspGlnAlaArgLeuSerAlaAspTyr 353
DB 1251 CAAGCCCAACCCACAGAGCGCGGAGAAAGTCCACAGAGCCCGGTTATCAGCTGAATAT 1310
QY 354 GluLysValGluLeuProLeuGluAspGlnValGlyAspLeuGluAlaSerSerGluGlu 373
DB 1311 GAGAAAGTTGAGCTGCCCTCAGAGGAGCAAGTCAGTGGCTCGCGGAGACCTTCTGAAGAG 1370
QY 374 LysCysAlaProLeuAlaThrGluValPheAspGluLysMetGluAlaHisGln---Glu 392
DB 1371 AAACCTGCTCCGTTGGCGACAGAGTGTGATGAGAAATAGAAAGTCCACCAAGAGAG 1430
QY 393 ValValAlaGluValHisValSerThrValGluLysThrGluGluGlnGlyGlyGly 412
DB 1431 GTTGTGCGCAAGTCCACGCTCAGCAGCGTGGAGAGAGAACCGAAGAGCAG----- 1481
QY 413 GlyLysAlaGluGlyGlyValValValGluGlyThrGlyLysSerLeuProGluLys 432
DB 1482 -----AAACCGAGGTGGAGAAACACAGAGGCTCTGTGCCAGCTGAAGAA 1526
QY 433 LeuAlaGluProGlnGluValProGlnGluAlaGluProAlaGluGluLeuMetLysSer 452
DB 1527 TTGTTTGGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1586
QY 453 ArgGluMetCysValSerGlyGlyAspHisThrGlnLeuThrAspLeuSerProGluGlu 472
DB 1587 AAAGAAACGTTGTTTCCGAGAGAGACCTTACACAGAGGAGTCACTCAGCTCTGATGAG 1646
QY 473 LysThrLeuProLysHisProGluGlyLysValSerGluValMetLeuSerSerGln 492
DB 1647 AAGTGTCTGTCAAAACCCCGGAGGCGTTGTGAGTGGAGTGAATGTGTCTATCAGAG 1706
QY 493 GluArgLysValGlnGlySerProLeuLysLysLeuPheSerSerSerGlyLeuLys 512
DB 1707 GAGAGATGAAGTGGAGGAGTCCACTAAAGAGCTTTTACCAGCACTGGCTTAAAA 1766
QY 513 LysLeuSerGlyLysLysGlnLysGlyLysArgGlyGlyGlyGlyGlyGlyGlyGly 532
DB 1767 AAGCTTCTGGAAGAAACAGAAAGGAGAAAGA--GGAGAGAGACGAGGAATCAGGG 1823
QY 533 GluTyrGlnHisLysThrGluSerProGluSerAlaAspGluLysGlyGlyGlyGly 552
DB 1824 GAGCACACTCAGGTTCCAGCCGATCTCCGAGCAGCAGAGGAGCAAGAGGCGGAGAGC 1883
QY 553 SerAlaSerSerProGluGluProGluGluThrCysLeuGluLysGlyProLeuGlu 572
DB 1884 TCTGCTCATCCCTGAGAGCCCGAGGAGATCAGCTGTCTGGAAAAAGGCTTAGCCGAG 1943
QY 573 AlaProGlnAspGlyGluAlaGluGluGlyThrThrSerAspGlyGlyLysLysArgGlu 592
DB 1944 GTCCAGCAGATGGGAGAGCTGAAGAGAGAGCTACTTCCGATGAGAGAAAAAAGAGAA 2003
QY 593 GlyLysThrProTrpAlaSerPheLysLysMetValThrProLysLysArgValArgArg 612
DB 2004 GGTGTCACTCCCTGGGCGATCATTTCAAAAGAGATGCTGACGCCCAAGAGCGTGTAGACGG 2063

[illegible]

QY	965	ArgThnGluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSerAla	984
Db	3144	GGTTCCGAAGAAGAACCGAAGCATCTGCTGCTGAAGAAGACACCAAAATGGTGTCAAGCA	3203
QY	985	ValSerGlnLeuThrAspSerProAspThrTrpGluGluValaThrProValGlnGluVal	1004
Db	3204	GTCTCCAGGTTTACCGACTCCCAAGACACCAAGAGAGGCGACTTCGGGTGAGAGGTG	3263
QY	1005	GluSerGlyValLeuAspThrTrpGluGluGluValArgGlnThrGlnAlaIleLeuGlnAla	1024
Db	3264	GAAAGTGCGGTAACCTGACATATGAAGAAGCAAGAAGCGGAGCTCAAGAGTCTCCAGGCA	3323
QY	1025	ValAlaAspIleValIleLysGluGluSerGlnValProAlaThr-----	1038
Db	3324	GTGGCAGAGAAAAGTGAAGAAGGAAATCCAGCTGGCCGTGCACCGGTGGCCAGAAAGATGTG	3383
QY	1039	---GlnThrValGlnArgThrGlnIleSerLysAlaLeuGluLysValGluGluValGluGlu	1057
Db	3384	CTTCAGCGCTGTGCAAGAGCA-----GAGGCAAAAAGACCAAGAAAG	3425
QY	1058	AspSerGluValLeuAlaSerGlnLysGluLysAspValMetProLysGlyProValGln	1077
Db	3426	CAGGCTGAAGCGTGGGTGTGAAGAAGAAGACGAGATGTATGTTGAAGTATGATGTCTAG	3485
QY	1078	GluAlaGluAlaGluIleLeuAlaGlnIleLysSerGluThrGlyGlnAlaThrProGluSer	1097
Db	3486	GAGCGAAAACCTGAGCCTTTTACACAAGGGAAGGTGTGGGCAACACCCCGAGAAGAC	3545
QY	1098	LeuGlu---ValProGluValThrAlaAspValAspHis-----ValAlaThrCys	1113
Db	3546	TTTGAAAAAGCTCTCTAAAGTCAACAGACATATGAGTCCAGTACCTTGTAAACCATTTGT	3605
QY	1114	GlnVal-----IleLysLeuGlnGlnLeu---MetGluGlnAlaValAla	1127
Db	3606	CAAGCCAAAACCTTACGTGGGGGTAAATATCAAGAGATGTATGAAACAGGCTATCCCC	3665
QY	1128	ProGluSerSerGlnThrLeuThrAspSerGluThrAsnGlySerThrProLeuAlaAsp	1147
Db	3666	CTTGACTGTGGTGAAGAACCTCTACAGACAGTGAAGCTGATGGAAGACCCCGGTAGCGAC	3725
QY	1148	SerAspThrAlaAspGlyThrTrpGlnGlnAspGluThrIleAspSerGlnAspSerLysAla	1167
Db	3726	TTTGAACGACCAAGGCACACACCCGAAAAGCAGAAATTTGGAAATCCATAGAGAGATGAG	3785
QY	1168	ThrAlaAlaValArgGlnSerGlnValThrGluGluGluAlaAlaThrAlaGlnLysGlu	1187
Db	3786	GTCCGATCTGTGTAACCAAGTACAGGGGGCACAAGAACAGAGCGAGTCTCTGCACAGAAAG	3845
QY	1188	GluProSerThrLeuProAsnAsnValProAlaGlnGluGluHisGlyGluGluProGly	1207
Db	3846	AGGCCTCAGCA---CCTTCAGATTTTGTGTTCCAGAGAAAGAAACATAAAGAACATCAAG	3902
QY	1208	---ArgAspValLeuGluProThrGlnGlnGluLeuThrAlaAlaAlaValProValLeu	1226
Db	3903	ATGGAAGCACTTGAAGCATATACATAAAGAGGTCTAGTGAAACTGTATCCATTCTG	3962
QY	1227	AlaIleThrGluValGluGlnGluGluLysGluValAspTrpLeuAspGlyGluLysValLys	1246
Db	3963	TCAAAAGCTAGGGGACTCA---GAGCGTACCACTATGCTGATGAGAAACCAAA	4016
QY	1247	-----GluGluGlnGluValPheValHisSerGly-----ProAsnSer	1259
Db	4017	GACGTACCATTTTTCGAAGGACCTTGAAGGGGTCTATGACACAGGATATAAAGTACGTGG	4076
QY	1260	GlnLysAlaAlaAspValThrIleTrpAspSerGluValMetGlyValAlaGlyCysGlnGlu	1279
Db	4077	GAAGAGCTCACTGAAGTTGCCCTTAAAGGTGAAGGACAGAAAGAAAGCTGAATGTATAAAG	4136
QY	1280	LysGluSerThrGluValGlnSer-----LeuSerLeuGluGluGlyLys	1294
Db	4137	GATGATGCTCTTGAACGCAAGATCAACGTAAGTCTCTTCATCCCGGTGAGAGAGAG	4196
QY	1295	MetGluThrAspValGluLysGluLysArgGluThrLysProGluGlnValSerGluGlu	1314

```
Db 4197 ATGGTAGTTCAGTCTGAAGGAGAGAAACAGACAGACGACCCACCTGTGAATGAAGAG 4256
Qy 1315 Gly---GluGlnGluThrAlaAlaProGluHisGluGlyThrTrpGlyProValLeu 1333
Db 4257 AAGCTTGAGCAGCAAAACAGCTGTACCGTATCTCAAGAGGTCAGTAAGCAGCTCCTCCAG 4316
Qy 1334 ThrLeuAspMetProSerSerGluArgGlyLysAlaLeuGlySerLeuGlyGlySer--- 1352
Db 4317 ACAGTGAATGTGCCATCATAGATGGGCAAGGAAGATGACAGCTTTTGGAGGAGGAGCCCT 4376
Qy 1353 ProSerLeuProAspGlnAspLysAlaGlyCysIleGluValGlnValGlnSerLeuAsp 1372
Db 4377 CTTCCCTGCTAGTCAAGAGGAGGAGTATGCACCAAAATTCAGTTTCAAGCTCTGAG 4436
Qy 1373 ThrThrValThrGlnThrAlaGluAlaValGluLysValIle-----GluThrValVal 1390
Db 4437 GCATCAATCACTCTAACAGCGGCTGCAGAGGAGGAGAAAGGCTTTAGGAGAACTGCCAAC 4496
Qy 1391 IleSerGluThrGlyGluSerProGluCysValGlyAlaHisLeuLeuProAlaGluLys 1410
Db 4497 ATTTTAGAAAACAGGTGAACCTTTGGAGCCTGCAGGTGCACATTTAGTTCTTGGAGAGAAA 4556
Qy 1411 SerSerAlaThrGlyGlyHisTrpThrLeuGlnHisAlaGluAspThrValProLeuGly 1430
Db 4557 TCCTCTGAAAAAATGAAGACTTTGCCGCTCATCCAGGGGGAAGATGCTGTGCCACAGGG 4616
Qy 1431 ProGluSerGlnAlaGluSerIleProIleIleValThrProAlaProGluSerThrLeu 1450
Db 4617 CCGGACTGTCCAGCAAAATCGACACCACTAGTATGATCTGCTACTACCAAGAAAGGCTTA 4676
Qy 1451 HisProAspLeuGlnGlyGluIleSerAlaSerGlnArgGluArgSerGluGluGluAsp 1470
Db 4677 AGTTCCGACCTGGAGAGAGAGAAAACACATCACTGAAGTGAAGTGCAGATGAAGTCGAT 4736
Qy 1471 LysProAspAlaGlyProAspAlaAspGlyLysGluSerThrAlaIleGluLysValLeu 1490
Db 4737 GAGCAGGTTGCTGCCAGGAGGTC-----AAAGTGAGTGTAGCAATTGAGGATTTA--- 4787
Qy 1491 LysAlaGluProGlu-----IleLeuGluLeuGluSerLysSerAsnLysIleValLeu 1508
Db 4788 -----GAGCCTGAAATGGGATTTTGGNACTTCAGACCAAAAGCAGTAACCTTGTCCAA 4841
Qy 1509 AsnValIleGlnThrAlaValAspGlnPheAlaArgThr---GluThrAlaProGluThr 1527
Db 4842 AACATATCCAGACAGCGGTTGACCACTTTGTACGTACAGAGAAACACGACCCGAAATG 4901
Qy 1528 HisAlaTrpAspSerGlnThrGlnValProAlaCysArgLeuAspSerArgGluProAsn 1547
Db 4902 TTGACGTCTGAGTTACAGACACAGCTCACGTGATATAAAGCTGACGCCAGGCGCTGGA 4961
Qy 1548 ArgCysTrpThrLys-----MetLysAspAla-LysMetLysHisProVa 1562
Db 4962 CAGAAACGAGAGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5021
Qy 1562 lProGlnProArgGlu-----AspLeuGlnValLeuThrValLe 1575
Db 5022 ACTTCAGCCAAAGAGGAGTCAAGCTCAACCGCAGTGGGCAAGCACATTCGTATATTTC 5081
Qy 1575 uGluAlaTrpAlaGlnProArgLysCysLeuProArgLeuGlnLeuLysAlaProVal 1594
Db 5082 AAAGACATGATGAGGCTCAGAAAGAACCATGCTGTTGAGGTAGAAGGTTCCACTG 5139
```

RESULT 5

```
US-09-880-107-3439
; Sequence 3439, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
```

```
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3439
; LENGTH: 6608
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 U81607
US-09-880-107-3439

Alignment Scores:
Pred. No.: 1,56e-292 Length: 6608
Score: 4300.50 Matches: 967
Percent Similarity: 68.81% Conservative: 189
Best Local Similarity: 57.56% Mismatches: 408
Query Match: 53.27% Indels: 116
DB: 10 Gaps: 36

US-09-902-432-4 (1-1596) x US-09-880-107-3439 (1-6608)

Qy 1 MetGlyValAlaGlySerSerThrGluGlnArgSerProGluGln---ProAlaGlySerAsp 19
Db 192 ATGGCGCCCGGAGGACTCCACCGAGCAGCGAGCGCGGAGCGCGCGCGCGGAGGAGTCC 251
Qy 20 ThrProSerGluLeuValLeuSerGlyHisGlyProAlaAlaGluAlaSer---GlyAla 38
Db 252 ACGCCGCTGAGCCGCGAGCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 311
Qy 39 AlaGlyAspProAlaAspAla-----AspProAlaThrLysLeuProGlnLysAsnGly 56
Db 312 ACCCGCGAGACCCCGCATCGCTGCTCGGACCCCGCCACCAAGCTCTTACAGAAAGATGGT 371
Qy 57 GlnLeuSerSerValAsnGlyValAlaGluGlnGlnGlyAspValHisValGlnGluAsn 76
Db 372 CAGCTGTCCACCATCAATGCGGTAGCTGAGCAAGATGAGCTCAGCTCCAGGAGGGTGAC 431
Qy 77 GlnGluGlyGln-----GluGlu 82
Db 432 CTAAATGGCCAGAAAGAGGAGCCCTGAACGGTCAAGGAGCCCTAAACAGCCAGGAGGAAGA 491
Qy 83 GluValValAspGluAspValGlyGlnArgGluSerGluAspValArgGluLysAspArg 102
Db 492 GAAGTCATTTGTCACGGAGGTTGGACAGAGAGACTCTGAAGATGTGAGCGAAAGAGACTCC 551
Qy 103 ValGluGluMetAlaAlaAsnSerThrAlaValGluAspIleThrLysAspGlyGlnGlu 122
Db 552 GATAAAGATGCTACTAAGTCAAGCGGTTGTTACGACATCAAGATGATGGCCAGGAG 611
Qy 123 GluThrSerGluIleIleGluGlnIleProAlaSerGluAsnAsnValGluGluMetVal 142
Db 612 GAGAACCGAAAT--ATCGAACAGATTCCTTCTTCAGAAAGCAATTTAGAAGAGCTAACA 668
Qy 143 GlnProAlaGluSerGlnAlaAsnAspValGlyPheLysLysValPheLysPheValGly 162
Db 669 CAACCCCACTGAGTCCCGGCTAATGATATTGGATTTTAAAGAGGTTTAAAGATTTTGTGGC 728
Qy 163 PheLysPheThrValLysLysAspLysAsnGluLysSerAspThrValGlnLeuLeuThr 182
Db 729 TTTAAATTTCACTGTGAAAAAGGATAAGACAGAGAGCTGACACTGTCCAGCTACTCACT 788
Qy 183 ValLysLysAspGluGlyGluGlyAlaGluAlaSerValGlyAlaGlyAspHisGlnGlu 202
Db 789 GTGAAGAAAGATGAAGGGGAGGAGCA-----GCAGGGGCTGGCGACCCAGGAC 839
Qy 203 ProSerValGluThrAlaValGlnGluSerAlaSerLysGluSerGluLeuLysGlnSer 222
```

Db 840 CCCAGCCTT-----GGGGCTGGAGAGACGATCTCCAAAGAAAGAAAGCCCAAAATCT 893
 QY ThrglulysglnugllyThrleuylsglnuglinserserThrglulileProleugln 242
 Db 223 ThrglulysglnugllyThrleuylsglnuglinserserThrglulileProleugln 242
 Db 894 ACAAGAAAGCCGAAAGAGACCTCTGAAGCGTGAGCAAGCCACGACGAATTTCTCCCA 953
 QY 243 AlaglserserProgluAlaIagluglulalylasrpglugllyglulysglnugl 262
 Db 954 GCCGAATCTGGCCAAAGCAGTGTG--GAGGAATGCAAGAGGAGGAGGAGGAGGAGGAGG 1010
 QY 263 LysgluProthrlyserProgluSerProserSerProvalAsmSerGluThrThrSer 282
 Db 1011 AAAGAACCTAGAAAGCTGCGAGAAATCTCGACTAGTCCCGTACCAAGTAAACAGATCA 1070
 QY 283 SerPheylslyPhePheThrHisglYTrpAlaglyTrpAlaglylsylsThrsPheyls 302
 Db 1071 ACCTTCAAAAAATTTCTTCACTCAAGTTGGCGCGGTGGCGCAAAAAAGCCAGTTTACGG 1130
 QY 303 LysSerlysgluAspAspLeugluThrAlaglylsylsArglysglnuglulalyls 322
 Db 1131 AAGCCGAAGAGAGATGAAGTGAAGTCTCAGAGAGAGGAGGAGGAGGAGGAGGAGGAGG 1190
 QY 323 ValAspGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 337
 Db 1191 GTAGACACAGAAAGAGACGAAAGGACAGAGTGTGCTCCGAGAGAACTGACCCGCTCCGAG 1250
 QY 338 -----GlnGluProAlaGluAspThrAspGlnAlaArgLeuSerAlaAspTyr 353
 Db 1251 CAAGCCACCCACAGAGAGCGGCGCAAGAAAGTGCCCAAGAGCCCGGTTATCACTGAATAT 1310
 QY 354 GlulysValGluLeuProleugluAspGlnValGlyAspLeugluLaserSerGluGlu 373
 Db 1311 GAGAAAGTGAAGTCCCTCCAGAGAGCAAGTCAAGTGTGCTCCGAGAGAACTTCTGAAGG 1370
 QY 374 LysCysAlaProleuAlaThrGluValPheAspGluLysMetGluAlaHisGln---Glu 392
 Db 1371 AAACCTGCTCGTGGCGACAGAGAGTGTGATGAGAAATGAAAGTCCACCAAGAGAG 1430
 QY 393 ValValAlaGluValAlaHisValSerThrValGluLysThrGluGluGluGluGluGly 412
 Db 1431 GTTGTGGCGGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1481
 QY 413 GlylulAlaGluGlyValValValValGluGlyThrGlyLysLeuProProGluLys 432
 Db 1482 -----AAAAAGAGGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1526
 QY 433 LeuAlaGluProGlnGluValProGlnGluAlaGluProAlaGluGluMetLysSer 452
 Db 1527 TTGTTGGAGATGATGACAGAACCTCAAGAGCCGAACCTGCAAGAGAGCTGGTGAAGCTC 1586
 QY 453 ArgGluMetCysValSerGlyGlyAspHisThrGlnLeuThrAspLeuSerProGluGlu 472
 Db 1587 AAAGAAACGTGTGTTCCGAGAGAGACCTTACACAGAGGAGCTGACTGCTGTATGAG 1646
 QY 473 LysThrLeuProLysHisProGluGlyValSerGluValGluMetLeuSerSerGln 492
 Db 1647 AAGGTGCTGTCCAAACCCCGCAAGAGCTGTGAGAGTGAAGTGAAGTGTGTATCATCAG 1706
 QY 493 GluArgGluLeuValGlnGlySerProleuLysLysLeuPheSerSerSerGlyLeuLys 512
 Db 1707 GAGAAATGAGGTGTCAGGAGGAGTCACTTAAGAGCTTTTAAACGACGCTGCTTAAAA 1766
 QY 513 LysLeuSerGlyLysLysGlnLysGlyLysArgGlyGlyGlyLysAspGluGluProGly 532
 Db 1767 AAGCTTTCTGGAAGAGAAACAGAAAGGAGAAAG---GAGAGAGAGAGACAGAGGATAGAGG 1823
 QY 533 GluYrGlnHisIleHisThrGluSerProGluSerAlaAspGluGluLysGlyLysSer 552
 Db 1824 GAGCACACTCAGGTTCACGCGATTTCTCCGACAGCCAGAGAGAGAGAGAGAGAGAGAGAG 1883
 QY 553 SerAlaSerSerProGluGluProGluGluThrThrCysLeuGluLysGlyLysProleuGlu 572
 Db 1884 TCTGCTCATCCCTGAGAGAGCCGAGAGAGATCACTGTGTGGAAGAGGAGGCTTAAAGCCGAG 1943

QY 573 AlaProGlnAspGlyGluAlaGluGluGlyThrThrSerAspGlyGluLysArgGlu 592
 Db 1944 GTGCAGCGAGATGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2003
 QY 593 GlyIleThrProTrpAlaSerPheLysLysMetValThrProLysLysArgValArgArg 612
 Db 2004 GGATCACTCCCTGGGCACTTCAATAAAGATGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2063
 QY 613 ProSerGluSerAspLysGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 632
 Db 2064 CTTTCGAAATGATTAAGAGATGAGCTGAGCAAGGTCAGAGCGCTTACTTGTCTTCC 2123
 QY 633 ThrAspSerThrValSerGluMetGlnAspGluValLysThrValGlyGluGluGluLys 652
 Db 2124 ACCGAGACACAGCGCTTGAATGCAAGAGAAATGAAGAGGAGCGTGGAGAGCCAAAG 2183
 QY 653 ProGluGluProLysLysArgValAspThrSerValSerTrpGluAlaLeuIleCysVal 672
 Db 2184 CCGAAGAAACCAAGCGAGGTGATCTCAGATCTCTGGAGAGCTTAAATTTGTGTG 2243
 QY 673 GlySerSerLysLysArgAlaArgLysAlaSerSerSerAspAspGluGlyLysProArg 692
 Db 2244 GGATCATCCAAAGAAAGAGCAAGAGAGAGTCTCTTCTGATGAGAGAGGAGGAGCAAAA 2303
 QY 693 ThrLeuGlyGlyAspSerHisArgAlaGluGluAlaSerLysAspLysGluAlaGlyThr 712
 Db 2304 GCAATGCGAGAGAGACCAACCAAGAGCTGATGAGCGCGGAGGAGGAGGAGGAGGAGGAG 2363
 QY 713 AspAlaValProAlaSerThrGlnGluGlnAspGlnAlaGlnGlySerSerSerProLys 732
 Db 2364 GACGGAGATCTTGTGTTGTTCCCAAGAAATATATCCAGGCGAGGAGGAGGAGGAGGAG 2423
 QY 733 ProAlaGlySerProSerGluGlyGlyGlyValSerThrTrpGluSerPheLysArgLeu 752
 Db 2424 CAAGCTGAGAACCTTACCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2483
 QY 753 ValThrProArgLysLysSerLysSerLysLeuGluGluGluGluGluGluGluGluGlu 769
 Db 2484 GTCAAGCCCAAGAAAAAATCAAGTCCAAAGCTGGAAGAGGAGGAGGAGGAGGAGGAGGAG 2543
 QY 770 ---SerSerValGluGlnLeuSerThrGluIleGluProSerArgGluGluSerThrVal 788
 Db 2544 GGGTGTGTGTAAGCAATTCCATCCACACACTGAACCGGTAAAGAAATCTCGGGTCTC 2603
 QY 789 SerIleLysPheIleProGlyArgArgLysLysArgAlaAspGlyLysGlnGluGln 808
 Db 2604 TCAATCAAGAGATTATCTCTGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2663
 QY 809 AlaThrValGluAspSerGlyProValGluIleAsnGluAspAspProAsnValProAla 828
 Db 2664 GCGCCGTGTGAAGAGCGAGGAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2723
 QY 829 ValValProleuSerGluTrpAspAlaValGluArgGlyLysMetGlu-----Ala 845
 Db 2724 GTGTGCTCTGTGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2783
 QY 846 GlnGluLeuThrGluLeuProGlnLeuGluGlyAlaValTrpValSerGluGluLeuSer 865
 Db 2784 CAAAAGCGCGAGAGACCCGAG 2843
 QY 866 LysThrLeuValHisThrValSerValAlaValIleAspGlyThrArgAlaValThrSer 885
 Db 2844 GAGAGTCAGTTAT 2903
 QY 886 ValGluGluArgSerProSerTrpIleSerAlaSerValThrGluProleuGlnHisThr 905
 Db 2904 ATTGAAGAAAGGTCTCTTCTTGTGATATCTTCAAGAGAGAGAGAGAGAGAGAGAGAG 2963
 QY 906 AlaGlyGluAlaMetProProValGluGluValThrGluLysAspIleIleAla---Glu 924
 Db 2964 GAGCTGAAGCGCACCTGTAACTAGAGAGGATATTGAAAGAGAGATTTGCAAGAGAA 3023

Db	1681	GAATAATGTTGAAGAAAGCTAGCTGAGAAATGTTGAAGAAAGCTAGCTGAGAAATGTTGAA	174
Qy	562	GluThrThrCysLeuGluLysGlyProLeuGluAlaProGlnAspGlyGluAlaGluGlu	581
Db	1741	GAAAGTGTAGCTGAAATGTTGAAGAAATCGTAGCTCCAACTGTTGAAGAAATCGTAGCT	1800
Qy	582	GlyThrThrSerAspGlyGluLysLysArgGluGlyIleThrProTrpAlaSerPheLys	601
Db	1801	CCAACCTGTT-----GAAGAAATGTTAGCTCCA-----AGTGTGTA	1836
Qy	602	LysMetValThrProLysLysArgValArgProSerGluSerAspGlyGluGluGlu	621
Db	1837	GAAAGTGTGCTCCAACT-----GTTGAAGAAAGCTAGTAGAGAAATGTTGAAGAAAGT	1890
Qy	622	Leu---GluLysValLysSerAlaThrLeuSerSerThrAspSerThrValSerGluMet	640
Db	1891	GTACTGAAAATGTTGAGAAAGCTAGCTGAAAATGTTGAAGAAAGCTAGCTGAAAAT	1950
Qy	641	GlnAspGluVal-----LysThrValGlyGluGluGlnLysProGluGluProLysArg	658
Db	1951	GTTGAAGAAATCGTAGCTCCAACTGTTGAAGAAATCGTAGCTCCA-----	1995
Qy	659	ArgValAspThrSerValSerTrpGluAlaLeuIleCysValGlySerSerLysLysArg	678
Db	1996	-----ACTGTTGAAGAAATGTTAGCTCCAACTGTTGTAGAAGGTGTG	2037
Qy	679	AlaArgLysAlaSerSerAspAspGluGlyGlyProArgThrLeuGlyGlyAspSer	698
Db	2038	GCTCCAAGTGTGAGAAAGCTAGTAGAGAAATGTTGAAGAAAGTGA-----	2085
Qy	699	HisArgAlaGluAlaSerLysAspLysGluAlaGlyThrAspAlaValProAlaSer	718
Db	2086	-----GCTGAAAATGTTGAAGAAAGCTAGCTGAAAATGTTGAAGAAAGT	2130
Qy	719	ThrGlnGluGlnAspGlnAlaGlnGlySerSerProGluProAlaGlySerProSer	738
Db	2131	GTAGCTGAAAATGTTGAAGAAAGCTAGCTGAAAATGTTGAAGAAATCGTAGCTCCA	2190
Qy	739	GluGlyGluGlyValSerThrTrpGluSerPheLysArgLeuValThrPro-----	755
Db	2191	GTTGAAGAAATCGTAGCT-----CCAACTGTTGAAGAAATGTTAGCTCCAAGTGTGTA	2244
Qy	756	-----ArgLysLysSerLysSerLysLeuGluGluLysAlaGluAspSerSerVal	772
Db	2245	GAAAGTGTGCTCCAAGTGTGAAGAAAGCTGTAGAAGAAATGTTGAAGAAAGTGTAGCT	2304
Qy	773	GluGlnLeuSerThrGluIleGluProSerArgGluGluSerTrpValSerIleLysLys	792
Db	2305	GAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGT-----	2346
Qy	793	PheIleProGlyArgArgLysArgAlaAspGlyLysGlnGluGlnAla-----Thr	810
Db	2347	-----GTAGCTGAAAATGTTGAAGAAAGTGTAGCTCCAAGTGTAGCTCCA	2382
Qy	811	ValGluAspSerGlyProValGluIleAsnGluAspProAsnValProAlaValAla	830
Db	2383	GTTGAAGAAATGTTAGCTCCAAGTGTGAAGAA-----AGTGTAGCT	2424
Qy	831	ProLeuSerGluTyrAsnAlaValGluArgGluLysMetGluAlaGlnGlyAsnThrGlu	850
Db	2425	CCAAAGTGTGAAGAAAGTGTGTGCTGAAAACGTTGCAACAAATTTTTCAGACAACTCT	2481
Qy	851	LeuProGlnLeuLeuGlyAlaValTyrValSerGluGluLeuSerLysThrLeuValHis	870
Db	2482	TTAAGTAATTTTATAGTGTATC-----GAACTGTAGAAATAAAGCACGATATATTAAT	2538
Qy	871	ThrValSerValAlaValIleAspGlyThrArgAlaValThrSerValGluGluArgSer	890
Db	2539	GAGATAGAA-----GAAGTAAAGAAAGT-----	2562
Qy	891	ProSerTrpIleSerAlaSerValThrGluProLeuGluHisThrAlaGlyGluAlaMet	910


```

QY 1196 ValProAlaGlnGluGluHisGlyGluProGlyArgAspValLeuGlu-----Pro 1213
Db 3619 -----GTTCAAGACAAATGATATGATGAAAGCTGTGAAAGATTTTAAAGATTAAGAAAAAT 3672
QY 1214 ThrGlnGlnLeuThrAlaAlaValProVal-----Leu 1226
Db 3673 ATGGAAGAGAGTTATGAAAGAGATGCTGTTGAAATTAATGACATTACTGCAAACTTAATT 3732
QY 1227 AlaIysThrGluValGlyGlnGluGluValAspTrpLeu---AspGlyGluVal 1245
Db 3733 GAAGAACTCAAGAGTTAATGAAAGTGAAGACGATTTAATAAGATATGGAATAAATTA 3792
QY 1246 LysGlu-----GluGlnGluValPheValHisSerGlyPro 1257
Db 3793 AAAGAAATTGAAAAAGCATTAATCAGAAAGATTCCTAAAGAAATATAGATCAAAAGATGAT 3852
QY 1258 AsnSerGlnLysAlaAlaAspValThrTyrAsp-----SerGluValMet 1272
Db 3853 ACATTAGAAAAAGTTATTTAGAGAGAACATGATATACACACGACGTTGGATGAAGTTGTA 3912
QY 1273 GlyValAlaGlyCysGlnGlu----- 1279
Db 3913 GAATTAAAAAGATGCGAAGAAAGACAGATCGAAAAAGTATCTGATTTAAAGATCTTGAA 3972
QY 1280 -----LysGluSerThrGluValGlnSerLeuSerLeuGlu 1292
Db 3973 GAAGATATATTAAGAAAGAAATCAAAAGAACTTGAAGTCAAAATTTTAAAGAT 4032
QY 1293 GlyGlu-----MetGluThrAspValGluLysGluValArgGluThrLysPro 1308
Db 4033 TATTAAGAAATTAAGAACTATTTGAAACAGATATTTAGAAAGAAAGAAATAGAAAAA 4092
QY 1309 GlnGlnValSerGluGlnGluGlnGlnGluThrAlaAlaProGluHisGluGlyThrTyr 1328
Db 4093 GATCATTTTGAAGAAATTCGAAAGAAAGCTGAAAGAAATTAAGATCTTGAGACGATATA 4152
QY 1329 GlyLysProValLeuThrLeuAspMetProSerSerGlu----- 1341
Db 4153 TTTAAAGAAAGTATCTTCATTAGAACTTGAAGAAAGAAATTTAGAGAACTACGAA 4212
QY 1341 ----- 1341
Db 4213 TTTAAAGAAAGGTAGAACATATATTAAGTGTGATCCGATATAAAGGTTTGAAGAA 4272
QY 1342 -----ArgGlyLysAlaLeuGlySerLeuGlyGlySer 1352
Db 4273 GATGATTTAGAAAGACTAGATGATTTAAAGAAAGTATTTAGACATGTTAAAGGAGAT 4332
QY 1353 ProSerLeuProAspGlnAspLysAlaGlyCysIleGluValGlnValGlnSerLeuAsp 1372
Db 4333 ATGGAATTAGGGGATATGATTAAG-----GAAAGTTTAAAGAA 4368
QY 1373 ThrThrValThrGlnThrAlaGluValAlaGluLysValIleGluThrValValIleSer 1392
Db 4369 GATGTAAACAACAACTTGAGAAAGAGTTGAATCC---TTAAAGAGTTTATCTAGT 4425
QY 1393 GlnThrGlyGluSerProGluCysValGlyAlaHis----- 1404
Db 4426 GCATTAGCGATGATGAAGAACAAATGAAACAAAGAAAAAAGCTCAAGAACTTAAGTTG 4485
QY 1405 -----LeuLeu-----ProAlaGluLysSerSerAlaThr 1414
Db 4486 GAAGAAAGTATTTATTAAGAAAGAGTAAAGAAAGAAACAAAGAAATTAACAAAGAAAG 4545
QY 1415 GlyGlyHisTrpThrLeuGlnHisAlaGlu-----AspThrValProLeuGlyProGlu 1432
Db 4546 AAAGTAAAGGTTTATTAAGTAAAGTAAAGAAACAAAGAGTGAATA---GTAGAAAGTTGAA 4602
QY 1433 SerGlnAlaGluSerIleProIleIleValThrProAlaProGluSerThrLeuHisPro 1452
Db 4603 ATGAAAGATGAAGATATA-----GAAAGAGTGTAGAAAGAA 4638

```

```

QY 1453 AspLeuGlnGlyGluIleSerAlaSerGlnArgGluThrGlySerGlnGluAspLysPro 1472
Db 4639 GATATAGAAAGAGATATGAAAGAAATTAAGTTAGACATATAGATGAAGATATGATGAA 4698
QY 1473 AspAlaGlyProAspAlaAspGlyLysGluSerThrAlaIleGluLysValLeuLysAla 1492
Db 4699 GATATAGGTGAAGCAAGAT-----GAAGTATATAGATTTAATAGTCCAAAAA 4746
QY 1493 GluProGluIleLeuGluLeuGluSerLysSerAsnLysIle 1506
Db 4747 GAGAAACCATTTGAAAGATTAAGCGAAAAAGAAAAATTA 4788

RESULT 7
US-09-742-096-1
; Sequence 1, Application US/09742096
; Patent No. US20020155441A1
; GENERAL INFORMATION:
; APPLICANT: DRUIHE, PIERRE
; APPLICANT: DAUBERSIES, PIERRE
; TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
; FILE REFERENCE: 200773USODIV
; CURRENT FILING DATE: US/09/742,096
; PRIOR APPLICATION NUMBER: 2000-12-22
; PRIOR FILING DATE: 1998-02-06
; PRIOR APPLICATION NUMBER: PCT/FR96/00894
; PRIOR FILING DATE: 1996-06-12
; PRIOR APPLICATION NUMBER: FR 95/07007
; PRIOR FILING DATE: 1995-06-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 6152
; TYPE: DNA
; ORGANISM: P. falciparum
US-09-742-096-1

Alignment Scores:
Pred. No.: 1,97e-23 Length: 6152
Score: 462.00 Matches: 323
Percent Similarity: 36.36% Conservative: 293
Best Local Similarity: 19.07% Mismatches: 646
Query Match: 5.72% Indels: 432
Gaps: 68

US-09-902-432-4 (1-1596) x US-09-742-096-1 (1-6152)
QY 53 GlnLysAsnGlyGlnLeuSerSerValAsnGlyValAlaGlnGlnGlyAspValHisVal 72
Db 527 CAAGTAAATGTCGATTAAGCTTACTTGT-----GAAAGAGTAAAGAAAAAATT 571
QY 73 GlnGlnGluAsnGlnGlnGlnGlu---GluGluValAlaAspGluAspValGlyGln 91
Db 572 CTGACCTTATTAAGAGAAAGAAATACATTAACTGAAGGTGATGATATTAATAATTATTA 631
QY 92 ArgGluSerGluAspValArgGluLys-----AspArgValGluGlu----- 105
Db 632 GAAGAAAGCCGAAGATATAAGAAATATCTTATTAAGTATATGAAGAACCAAAAGAA 691
QY 106 -----MetalAlaAsnSerThrAlaValGluAspIle 116
Db 692 AATATATTATGACAAATTTATTAATATATTTGACAAAAATTCGAAAAACACAGAAAGTGA 751
QY 117 ThrLysAspGlyGln-----GluGluThrSerGluIleIleGluGlnIleProAla 133
Db 752 TCAGAAAAATGTCAGATCGATGATGACCTTTTAATGAAATTTAAATAGGTGTAGATGTT 811
QY 134 SerGluAsnAsnValGluGluLysValGlnProAlaGluSerGlnAlaAsnAspValGly 153
Db 812 AATGGAAGATGAAGAAATATTTTGGAG-----GAAAGTCAAGTTATATGACGATATT 865
QY 154 PheLysLysValPheLysPheValGlyPheLysPheThrValLysLysAspLysAsnGlu 173

```



```
Db 4730 GAAGAAGTATTATTATTAAGAAGAGGTTTAAAGAACCAACCAAGAAAAATAACAAAAAAG 4789
Qy 1415 GlyGlyHisTrpThrLeuGlnHisAlaGlu-----AspThrValProLeuGlyProGlu 1432
Db 4790 AAAGTAAAGTTTGTATATTAGGATTAAGAACCAACCAAGATGAATA---GTAGAAGTTGAA 4846
Qy 1433 SerGlnAlaGluSerIleProIleValThrProAlaProGluSerThrLeuHisPro 1452
Db 4847 ATGAAGATGAAGATATA-----GAAGAAGATGTAGAAGAA 4882
Qy 1453 AspLeuGlnGlyGluIleSerAlaSerGlnArgGluArgSerGluGluGluAspLysPro 1472
Db 4883 GATATAGAACAAGATATAGAAGATAAAGTTGAAGATATAGATAGATATAGATGAA 4942
Qy 1473 AspAlaGlyProAspAlaAspGlyLysGluSerThrAlaIleGluLysValLeuLysAla 1492
Db 4943 GATATAGTGAAGACAAAGAT-----GAAGTTATAGATTTAATAGTCCAAAAA 4990
Qy 1493 GluProGluIleLeuGluLeuGluSerLysSerAsnLysIle 1506
Db 4991 GAGAAACGCATTGAAAGGTTTAAAGCGAAAGAAAAAATTA 5032
```

RESULT 8

```
US-09-864-761-17726
; Sequence 17726, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeonica-X-1
; CURRENT APPLICATION NUMBER: US/09/864, 761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
```

```
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 17726
; LENGTH: 7997
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL034555.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 10
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 8.9
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.8
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 14
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 7.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 9.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 7.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 9.3
; OTHER INFORMATION: EXPRESSED IN BT100, SIGNAL = 7.7
; OTHER INFORMATION: EXPRESSED IN HBL474, SIGNAL = 12
; OTHER INFORMATION: EST HUMAN HIT: AUL17052.1, EVALUE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: P08640, EVALUE 4.00e-10
; OTHER INFORMATION: NT HIT: gi11427711, EVALUE 0.00e+00
US-09-864-761-17726
Alignment Scores:
Pred. No.: 3 69e-21 Length: 7997
Score: 431.50 Matches: 373
Percent Similarity: 34.27% Conservative: 271
Best Local Similarity: 19.85% Mismatches: 759
Query Match: 5.34% Indels: 476
DB: 10 Gaps: 83
```

US-09-902-432-4 (1-1596) x US-09-864-761-17726 (1-7997)

```
Qy 31 ProAlaAlaGluAlaSerGlyAlaAlaGly-AspProAlaAspAlaAspProAla----- 48
Db 886 CCTGGCAGCAAACTTCAGGCTCTCTCAGCGAGCGCTGCAAAATCTGACTGTCTAACT 945
Qy 49 -----ThrLysLeuProGlnLysAsnGlyGlnLeuSerValAsnGlyVa 64
Db 946 GGAATCAGTTAGTAATCAAAAGTACAAAGGAAAGGGCTTTCAAGCCATCTTCAAGTGT 1005
Qy 64 lAlaGluGlnGlyAspVal-----HisValGlnGluGlnGlnGluGlyG1 80
Db 1006 GGAGAAGGAGGCGGCTTAAGCCAGGAGCACCTCAAGCCTTGAG---CAGCCTGCAGA 1062
Qy 80 nGluGluGluValValAspGluAspValGlyGlnArgGluSerGluAspValArgGluLy 100
Db 1063 TGGGGTAAGTGTGTGAT-----1081
Qy 100 sAspArgValGluGluMetAlaAlaAsnSerThrAlaValGluAspIleThrLysAspG1 120
Db 1082 -----CTGGAGAAGCTGGAAGCCAGGAAAGCGCTTTGCAGATTTCAATTTAAAGC 1134
Qy 120 yGlnGluGluThrSerGluIleIleGluGlnIleProAlaSerGluAsn---AsnValG1 139
Db 1135 AGAAAAAGCAAAACCAGAGGTCAGAAAGAGCAGTCCAGAGATGGAGGATGCTCGCGTGT 1194
Qy 139 uGluMetValGlnProAlaGluSerGlnAlaAsnAspValGlyPheLysLysValPhe-- 158
Db 1195 TTCAAAAAGACGACCT-----GACGTGCTCTCTAGAGAGGTCTATTCT 1236
Qy 159 -----LysPheValGlyPheLysPheThrValLysLysAs 170
Db 1237 GCTGAGGAGGAGAGGCTGMAAAGAAAGCTGTGAGGAAAGAAATTTCTTAAAGAGATC 1296
Qy 170 pLysAsnGluLysSerAspThrValGlnLeuLeuThrValLysLysAspGluGlyGluG1 190
Db 1297 TAAAAAAATCAAACTGGACAGACTTAATCTGTTCAGCCCAAGACTGTCCAGGAGCT 1356
Qy 190 yAlaGluAlaSerValGlyAlaGlyAspHisGlnGluProSerValGluThrAlaValG1 210
Db 1357 TGCCAGTATTCTGTGGGTCTGGCTCAAGGCCAGCTCAGACCTCAAGAGACTGGG 1416
Qy 210 yGluSerAlaSerLysGluSerGluLeuLysGlnSerThrGluLysGlnGluGlyThrLe 230
```

```

Db 1417 AGAAGTAGCAGGT-----GAATCTGGGAAATCAAGAA----- 1450
Qy 230 uLysGlnGlnSerSerThrGluLePro-----LeuGlnAl 243
Db 1451 -----GTCCATCAAAAAGAACCCCATCTCCCAAAACACAGCTCAACAGCTGAGGT 1503
Qy 243 AGluserSerGlnAlaAlaGluGluGlu-----AlaLysAspGluGluGluGluGluGlu 261
Db 1504 ATTAGATGATTCAGAGCAGAGAGAGAGAACGTTAGAGAAACATATTGAGCTTCCTGTA 1563
Qy 261 nGluLysGluProThrLysSerProGluSerProGluSerProValAsnSerGluThr-- 280
Db 1564 TGAAACACCTGAAACGTAATCAAGGCCAAGAGAAATCACTTACGTAATATCTGAAGAAA 1623
Qy 281 -----ThrsSerPheLysLysPhePheThrHisGlyTTrpAl 293
Db 1624 AATTGGCATTGACATTCGATCAACGAGAGTTACCGAAACAAATGAGACAGAGTCTGAG 1683
Qy 293 AGlYTrpAlGlyLysLysThrSerPheLysLysSerLys-----GluAspAs 308
Db 1684 GAAACAGCAGATGAGAAATGCAAAATAGCAAGCTGAGAAATTGGCAGTCTCAAAAAGAA 1743
Qy 308 nLeuGluThrAlaGluLysArg----- 315
Db 1744 TGTAGATGAAATGTGAAGAGCTAGCCTCTTCAAGAGTAGCCCTTCAAGATGT 1803
Qy 316 -----LysGluGlnGluAlaGluLysValAsp 324
Db 1804 CACTGATGACTCTCTCTCCAGCAAAAGAAAGAGATGATCATGTGATTTGATATCTG 1863
Qy 325 -----GluGluGluLysGluLysThrGluProAlaSerGluGlnGluProAl 341
Db 1864 CACCAAGCAGAGAACGGAATTACAGAAATTCACGCCCAATACAGGAAGATTCTGAAAGAC 1923
Qy 341 AGluserThrAspGlnAlaArgLysSerAlaAspLysValGluLeuProLeuGlu 361
Db 1924 TGGTGTTCTCCAGCTGCCAGATGCTTCCTTCATGAAAGAT-----GA 1968
Qy 361 uAspGlnValGlyAspLeuGluAlaSerSer-----GluGluLys 374
Db 1969 GGATCCATCAGGTCTCCCTAGGCTACTGACATGAAAGAGTCTCTAAAGATGATGAAA 2028
Qy 374 sCysAlaPro-----LeuAlaThrGluValPheAs 384
Db 2028 ACTCTCTCCCTTATTACATTAACAGTCAAGAGAGAGTCTTTAAATTAATCTTATGA 2088
Qy 384 nGluLysMetGluAlaHisGlnGluValAlaAlaGluValHisValSerThrValGluLys 404
Db 2089 TTCTAGCAGAGAGAGAACAGATGGCAGATATGGCCAAATTAATACTATCTGTCTTGAAATTC 2148
Qy 404 sThrGlnGluGlnGluGlnGlyGlyGlyGluAlaGlu-----GlyGlyValValVa 421
Db 2149 TGAAGATGAACATAATCGTTGGAGCTCTCGAGATGAAGACAGATGCGGACGATTTGATGT 2208
Qy 421 iGluGluThrGlyGluSerLeuProGluLysLeu-----Al 434
Db 2209 GAGTTTCCCAACAGCATATTAAGAGAGATAGCTTCGAAAAAGAGTCTGTACGAGATCT 2268
Qy 434 AGluserProGlnGluValProGlnGluAlaGluProAlaGluGlu----- 448
Db 2269 GGAACCTGGGTGAGGTGCTTCTGATTCTGACGAAGATGGTGAACACAAATCCCACTCACCC 2328
Qy 449 -----LeuMetLysSerArgLysLys 455
Db 2329 CAGAGCCTTCGATTTATGAAGTTCTCGATTGTCTTTTATTATGAGGACAGAGAAAGA 2388
Qy 455 tCysValSerGlyGlyAspHisThrGlnLeuThrAspLeuSerProGlu----- 471
Db 2389 CAAGCTACGTGAGCAGATGAAGAACTCTCTAGTCTTTAGAAAGAACAAATTTTACTC 2448
Qy 472 -----GluLysThrLeuProLysHisProGluGluGlyLeValSerGluValGluMe 488

```

```

Db 2449 TTTTCATTGGATTAAGACATCAACAGACACTAAAGCTTTGCTTGAAGAGTAATTC 2508
Qy 488 tLeuSerSer----- 491
Db 2509 CCTCTCTTCACTCGTGAAGAAATTGGCTTTTCTTGATTGGAGCTCCCGATTTGCAAA 2568
Qy 492 -----GlnGluArgIleLysValGlnGlySerPro----- 501
Db 2569 TTTTGAACAACAAGATTAAGAAAGTTGACTCTCTCCACAGACTTATTCATCTCTG 2628
Qy 502 -----LeuLysLysLeuPheSerSerSerGlyLeuLysLysLeuSerGlyLysGlu 519
Db 2629 GTACATGAAAAAGAAAGAAATTAGAGCTGATTCAGAGGAAATAGATGATTAAGAAAGA 2688
Qy 519 nLysGlyLysArgGlyGlyGlyGlyAspGluGluProGlyGluThrGlnHisIleHisTh 539
Db 2689 GGACCATTA-----GAAAGAGAGCAGAGAGAGCAGAGATGTTGTTCTTC 2733
Qy 539 rGluSerProGluSerAlaAspGluGlnLysGlyLysLeuSerSerAlaSerSerProGlu 559
Db 2734 TCGTTTTTTACACAGCTCAATCTTTGAACAAAGATTCACAGGATTCAGCATTCAGAGAG 2793
Qy 559 uProGluGluThrThrCysLeuGluLysGlyProLeuGluAlaProGlnAspGlyGluAl 579
Db 2794 AAAAGAGAGAGATTCGACTTCATTTCTGTAGAGATCTATGGAGACAGACATCTGAAGG 2853
Qy 579 AGlGluGluGlyThrThrSerAspGlyLysLysArgGluGluGlyLysLysThrProTTrpAlaSe 599
Db 2854 AGCAAAACAGCACACT-----GATTCATTCAMAGAACAGTAGTCTGTTCCATAG 2904
Qy 599 r-----PheLysLysMetVal-----ThrProLysLysArgValaLysArgProSe 614
Db 2905 CAGATTTTGAGAGCTCACACGAGTCAACAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 2964
Qy 614 rGluSerAspLysGluGluGluGluGluGluLysValLysSerAlaThrLeuSerSerThrAs 634
Db 2965 AGAGGTTAGAGAAACAGAGAGATACAGAGATCATCCAAAGCCCAAGATGCTGCTCGTA 3024
Qy 634 pSerThrValSerGluMetGln-----AspGluValLysThrVa 647
Db 3025 GAATTAAGATTCAGAACTGAAACATCCACCTTCCTGGGCGCTCCAAAGTGCACAGCTGT 3084
Qy 647 iGlyGluGluGluLysProGluGluProLysArgArgValAspThrSerValSerTrpGlu 667
Db 3085 AACTCTAGATCAGCCCACTACAGCACTAGAGAGACCACTGGTGCAAAAACGAGTAGAGGC 3144
Qy 667 uAlaLeuIleCysValGlySerSerLysLysArgAlaArgLysAlaSerSerSerAspAs 687
Db 3145 GCCTTTGTA-----ACGAGAGAGAGAGCTGTGAGCCAGCTTACCTCTCGAGAGA 3195
Qy 687 p-----GluGlyGlyProArgThrLeuGlyLysAspSerHisArgAlaGlu 702
Db 3196 AGCAAAAGCTGCATGTGAACCTGCTGCCCCCTGT-----GAACACTGGA 3243
Qy 702 uGlu-----AlaSerLysAspLysGluAlaGlyThrAspAlaValPr 716
Db 3244 ACAAGTAGACCTGCCCCCAAGAGAGAGACCCCGATTAAGAGAGCTGCC-----ATGATGCC 3297
Qy 716 oAlaSerThrGlnGluGlnAspGlnAlaGlnGlySerSerSerProGluPro----- 733
Db 3298 TCGGGGTGTGAGAA-----GTTCACTAGGTGACCAAGCCCTTATCT 3342
Qy 734 -----AlaGlySerProSerGluGlyGlyValSerThrTrpGlu 747
Db 3343 GGATGCAAGCCTCCACAGCTCCGGGGCGTCGTTTCCAGGAGAGAGAACAGTAGATCC 3402
Qy 748 -----SerPheLysArgLeuValThrProArgLysLysSerLysSerLysLeuGlu 764
Db 3403 AGAGCCTGACATACCCAGCCCACTTCAAAACCACTCAGAAAGTCTGAGAGAACCAATTA 3462
Qy 764 uGluLysAlaGlu-----AspSerSerValGlu----- 773
Db 3463 GCCAAAGGCCGAAAGCCAGAGCCCACTGCAATGCTGAGCTGATGCAAACCAAGAAAGC 3522

```

QY 774 -GlnLeuSerThrGluLeuGluProSerArgGluGluSerTrpValSerIleLysLysPhe 793
DB 3523 CGAAGCTGCTCTGAGTCTCAGCCCGAGCTTCTGAAGAT---TTAGAGTGTGATCTCC 3579
QY 793 eileProGlyArgArgLysLysArgAlaAaspGlyLysGlnGluGlnAlaThrValGluAs 813
DB 3580 AGTTGTGCAAGAGTAAAGCCAAACAAAGCAAGCGTTCAAAGACCCCTGTTCAG-- 3637
QY 813 pSerGlyProValGluLeuLeuAsnGluAaspProAsnValProAlaValProLeuSe 833
DB 3638 -GCAGCTGCAGTGTGATCTGTGGAG-----AAGCCCGTCAAGGAAGAG 3681
QY 833 xGluTyrAsnAlaValGluArgGluLysMetGluAlaGlnGlyAsnThr-----GluLe 851
DB 3682 TGAG-----AGGATAGCCGGGAAACAACTCAAGCGGTCCAATTCTCTCGGGGAGAAGC 3735
QY 851 uProGlnLeuGlyAlaValTyrValSerGluGluLeuSerLysThr----- 867
DB 3736 ACAGAGCTTTTGAATTGAATGAGATGGAGGAGAGAGATACAGGACTGCTTCTAATAAA 3795
QY 868 -----LeuValHisThrValSerValAlaValIleAaspGlyThrArgAla-- 882
DB 3796 CTCTGCTGCAGACCTTGAACATCCGAAACCAAGTTTGCCTCTCAGCCGAAACAGGCGCG 3855
QY 883 -----ValThrSerVal-----GluGluArgSerProSerTrpIle 894
DB 3856 GAATGTAAAGAGGCTCTATGCAACCATGGGTGACCATGAAACCGCTCTCTCT----- 3907
QY 894 eSerAlaSerValThrGluProLeuGluHisThrAlaGlyGluAlaMetProProValG1 914
DB 3908 -----GTCAAGAGCCGTTGAGCAA-----CC 3930
QY 914 uGluValThrGluLysAspIleAlaGlu-----GluThrProValLeuThrGlnTh 932
DB 3931 AAGAGTGACCAGAAAGAGATTTGGAGCGAGAGCTTCAGGAGGCTGCAGCGGTTCCACCAC 3990
QY 932 r-----LeuProGluGlyLysAspAlaHisAaspMetValThrSerG1 947
DB 3991 CCCTCGAGGGGAGGCTCCAAAGACACCGCGCGAGCGATGAAGAGGAGGAGAACGA 4050
QY 947 uValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeuArgThrG1 967
DB 4051 GGCAAGAGAACCTGCAGAAACACTCAAGCCACTGAGGGATGGCGTCCCAAGGTCCCA 4110
QY 967 uGluValThrGluAlaSerGlyAlaGlu-----GluThrThrAspMetValSe 983
DB 4111 GAAACTGCAGCTGCTGTGACCCCAAGGAAAGGAAAGGAAATGAACCGAAGTGA 4170
QY 983 rAlaValSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnG1 1003
DB 4171 TGCTACA-----CGTCTGAGGCCACCACTGAGGTGGCGCCCAAAATAGG 4215
QY 1003 uValGluSerGlyValLeuAspThrGluGluGluGluArgGlnThrGlnAlaIleLeuG1 1023
DB 4216 CGTGAAGAGAGCTCCATGAAACCAAGCGTCTGAGAGGAGGAGGAGTGAACAGAA 4275
QY 1023 nAlaValAlaAspLysValLysGluGluSerGlnValProAlaThrGlnThrValGlnAr 1043
DB 4276 ACGTGACAGAAAGATCTGCGCACAGACAAACCCCTGAAACCGCCCTGTGAA-- 4333
QY 1043 gThrGlySerLysAlaLeuGluLysValGluGluValGluGluAaspSer----- 1059
DB 4334 -----GTTGTAGAGAAACCCGCGCCCTGAAACAACTCCAATCAAGAG 4380
QY 1060 -----GluValLeuAlaSerGluLysGluLysAs 1069
DB 4381 AGGAAGATCTCGAAATCTCAGGTTAGCAGTGGCAAAATCTGCAAGTCTGAAATAATGCGA 4440
QY 1069 p-----ValMetProLysGlyProValGlnGlnAlaGlyAlaGluHisLeuAlaGlnG1 1087
DB 4441 TGCTGTGTCTCCAGGGGGCTGCAGCACAGGCGAGGAGG----- 4486

QY 1087 ySerGluThrGly-----GlnAlaThrProGluSerLeuGluValProGluValThrAl 1105
DB 4487 ----GAATCTGGGGTGGTGCACTCTCCCTTGAGAAAGTAGAGTCCCAAGAGGAGGA 4542
QY 1105 aAspValAspHisValAlaThrCysGlnValIleLysLeuGlnGlnLeuMetGluGlnAl 1125
DB 4543 TGGTTTATCA-----TCCAGTGTGAAAAGTGTATCC 4572
QY 1125 aValAlaProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrProLe 1145
DB 4573 AGTTGATCCAGACAGAACCCAGAGAAAGAGAGCTGTCTGCTCTGGGCGTCCCCAGA 4632
QY 1145 uAlaAspSerAspThrAlaAspGlyThrGlnGlnAspGluThrIleAspSer----- 1162
DB 4633 AGCC-----ACCAGTTAGCCAAAGCAGATGGAGCTGGAGCGCGTGGAAACACATCGCAAA 4689
QY 1163 ---GlnAspSerLysAlaThrAlaAlaValArgGlnSer----- 1174
DB 4690 GCTCGCTGAGCGCTCTGCTCTGCTGCTATAGGCAGATGACACAGAGGCGCTTGCCTCC 4749
QY 1175 -----GlnValThrGluGluGluAlaAlaThrAlaGlnLys 1186
DB 4750 AGAGCAGGAGCAAGCTGCACACCAAGCAAGTGAACAGAGCTGCTCGCGCCATCGG 4809
QY 1186 s-----GluGluProSerThrLeuProAsnAsnValProAlaGlnGluHis-- 1202
DB 4810 CTCCATCATCAATGACATTCTCGGGAGCGAGAAACTTCCAGCACCTCCACCTTATCC 4869
QY 1203 -GlyGluGlu-----ProGlyArgAspValLeuGluProThrG1 1215
DB 4870 TGAGAGATCCACAGACAGATCTGCAACCCCGCAGGTGCACAGCGCTGACGCTTCTGA 4929
QY 1215 nGlnGluLeu---ThrAlaAlaAlaValProValLeuAlaLysThrGluValGlnG1 1234
DB 4930 GGAAGGAATGAGACAGATGAGGCTGTATCTGTCATCTCTGGAACCTGAGGCTGCTACA 4989
QY 1234 uGlyGluValAspTrpLeuAspGlyGluLysValLysGluGluGlnGluValPheValHi 1254
DB 4990 ATCTTCTAGCGCTCCAGTCAATGCTCTCTGACCCCTCA----- 5026
QY 1254 sSerGlyProAsnSerGlnLysAlaAlaAspValThrAspSerGluValMetGlyVa 1274
DB 5027 -GCCGCCCAACAGATACCAAGGAAGCCAGAGGAATAGCAGTGAACCTCACACTCAGT 5085
QY 1274 lAlaGlyCysGlnGluLysGluSerThrGluValGlnSerLeuSerLeuGluGluG1 1294
DB 5086 GCCAGAAGCAAGAGGTCTAAAGAGTGAAGTCACTCTTGTTCGGAAAGACAAAGGCG 5145
QY 1294 uMetGluThrAspValGluLysGluLysArgGluThrLysProGluGlnValSerGluG1 1314
DB 5146 CCAGAAAAACAACCCGATCAGCCGCAAGCGAAACACAAACAGAAAGTGTGCT----- 5200
QY 1314 uGlyGluGlnGluThrAlaAlaProGluHisGluGlyThrTyrGlyLysProValLeuTh 1334
DB 5201 ----CTGTAGAGCGATGCTCCGTAATCCAACCAAGCTCAAGGT----- 5242
QY 1334 rLeuAspMetProSerSerGluArgGlyLysAlaLeuLysSerLeuGlyGlySerProSe 1354
DB 5243 ----GAGAGTCTGTGCTGCAATAGGCGGACACAGTA-----CAGCACCCCGA 5286
QY 1354 rLeuProAspGlnAspLysAlaGlyCysIleGluValGlnValGlnSerLeuAspThrTh 1374
DB 5287 AGCCCAACAGGAGAAAG-----CAGAGTGAAGAAACCCCATTCAC 5328
QY 1374 rValThrGlnThrAla-----GluAlaValGluLysValIleG1 1387
DB 5329 TCCTCTCAGTCACTGACTTCTGACCTAGCAAGATTCCTCCACAGAGAAATTCGTCCTCA 5388
QY 1387 uThrValValIleSerGlyThrGlyGluSerProGluCysValGlyAlaHisLeu----- 1405
DB 5389 AGAAATCAGTGTGAGAAAGGACTCCACCAAGAGCATCTGTGCCCCAGACCTTCCCCC 5448
QY 1406 -----LeuProAlaGluLysSerSerAlaThrGlyGlyHisTrpThrLeuGlnHisAl 1423

```
Db 5449 ACCTCCCGACGACGCGGTGATGAGAGCTTCAAGCCAGGTTGAGGTGATTCAT 5508
      |||||
Qy 1423 agluaspThrvAlProleuglyProgluSerInlaIagluSerIleProIleValTh 1443
      |||||
Db 5509 CATTAAGTGAACCGGTATACCCACCCAGCGAT---CCAAGATCCCATACCCAGACT 5565
      |||||
Qy 1443 rProAla-----ProGluse 1448
      |||||
Db 5566 GCGTCTGTACTGACGACGAAGCTCTACCTCTGCTGCGCTGTGGGGAGATCCACACCA 5625
      |||||
Qy 1448 rThruenHspProaspIeuglnlygluIleSerAlaserGlnargGluArgSerIuGl 1468
      |||||
Db 5626 GAGCCCCCTTACTAAGGTACAGAGTGCATCAAGGCGAGGAGGCCAGGCGCTAGTC 5685
      |||||
Qy 1468 uGluaspIySPProaspAlaGlyProaspAlaaspIySglu-----SerTh 1484
      |||||
Db 5686 TACTGCATCTCCAGCTCTCCCGACAGACAAAGGCTCTGTATGTGACACCAAGCTCCAG 5745
      |||||
Qy 1484 rAlaIleGluIyValIleuIyAlaIgluProgluIleuIyGluIleuIySerIySerz 1504
      |||||
Db 5746 CACCTGAGGAAGATTCTC---ATGAGACCCCAAGTATGTGTGCGCACCAAGGTCACTTC 5802
      |||||
Qy 1504 nLysIleValIleuIyValIleGlnThrvAlaIaspGlnPheAlaArgThrvGluThrvAl 1524
      |||||
Db 5803 CACAAGTGTACACACAGCCATTCAGAGCTGTC-----AGTCTGC 5844
      |||||
Qy 1524 aProgluThrvAlaIyAspSerGlnThrvAlaIProAlaCySArgIeAspSerAr 1544
      |||||
Db 5845 CCCTTGCCACT-----GAGGCCCGCGCCCGCCAGCTGACTTAA 5886
      |||||
Qy 1544 gGluProaspArgCySTPThrvIySmerIySaspAlaIySmerIySHIspProAlProgl 1564
      |||||
Db 5887 AAAGCT-----TTAGAGAAAAAAGACGACCTCCAGTGCACAA 5925
      |||||
Qy 1564 nProargIuIyaspIeuglnValIleuThrvAlaIgluIyATPAlaIgluProargIyCy 1584
      |||||
Db 5926 C---AACTGTGAGATTCAGACCTCGAGGTGCTGTAGCT---GCTGACAGAAAGGT 5979
      |||||
Qy 1584 sIeuProargIeuglnIleuIyS-----AlaProValSer 1595
      |||||
Db 5980 GGCTCCAGTCACTTGCTCCCAAAATTACTCTGTATTAGCCGAGTCCCTGTACG 6034
      |||||

RESULT 9
US-09-917-800A-1505
; Sequence 1505, Application US/09917800A
; Patent No. US20020119462A1
; GENERAL INFORMATION:
; APPLICANT: Mendick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Casle, Arthur
; APPLICANT: Elashoff, Michael
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5038-US
; CURRENT APPLICATION NUMBER: US/09/917,800A
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/222,040
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: US 60/222,880
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: US 60/290,029
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/290,645
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: US 60/292,336
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/295,798
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/297,457
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,884
```

```
      ; PRIOR FILING DATE: 2001-06-19
      ; PRIOR APPLICATION NUMBER: US 60/303,459
      ; PRIOR FILING DATE: 2001-07-09
      ; NUMBER OF SEQ ID NOS: 1740
      ; SOFTWARE: PatentIn Ver. 2.1
      ; SEQ ID NO 1505
      ; LENGTH: 15231
      ; TYPE: DNA
      ; ORGANISM: Rattus norvegicus
      ; FEATURE:
      ; OTHER INFORMATION: Genbank Accession No. US20020119462A1 X59601
      ; US-09-917-800A-1505

Alignment Scores:
Pred. No.: 1,63e-20 Length: 15231
Score: 427.00 Matches: 366
Percent Similarity: 33.95% Conservative: 258
Best Local Similarity: 19.91% Mismatches: 745
Query Match: 5.29% Indels: 471
Db: 10 Gaps: 70

US-09-902-432-4 (1-1596) x US-09-917-800A-1505 (1-15231)
Qy 3 AlaIySerSerThrvGluIuIyArgSerProgluIuIyProAlaIySerAspThrvProSer 22
      |||||
Db 4467 GCTGGCTAGACACAGCGCGGACGA-----GAGCGCGAGCGCTGCGC 4508
      |||||
Qy 23 GluIeuValIeuSerGlyHieGlyProAlaIaIagluIaSerGlyAlaIaGlyAspPro 42
      |||||
Db 4509 CGA-----GTTGAGAGCGCGCGCTGTGAGAACAGCGCGGCGTGTGAGAGCCCA 4556
      |||||
Qy 43 -AlaaspAlaaspProAlaIuIySleuProgluIuIySasnIyGlnIeuSerSerValAs 62
      |||||
Db 4557 TGCCGAGCGCAAGGACAGAGCGCGGCTGAGAGCAGCAAGTGCAGCGCGCATCCAGGA 4616
      |||||
Qy 62 nGlyValAla-----GluGlnIyAspValHieValGlnIuIySasnIyGlnIyG 80
      |||||
Db 4617 GAGGTGAGCGCGCGCGAGAGCGCGCGGTGAGACGACAGCAAGAGCGAGCATCA 4676
      |||||
Qy 80 nGluIuIy-----GluValValAs 86
      |||||
Db 4677 AGAGAGCTGCAGCATCTCGCGCAAGCTCAGAGGACAGATCCAGGCCAAGGCCACGA 4736
      |||||
Qy 86 pGluaspValIyGlnIyArgIuIySerGluaspValIyArgIuIySasnIyGlnIyGluIu 106
      |||||
Db 4737 GGTGAGCGCTGCAGCGCGCAGCGCATCGCATTCAGAGAGAGATCCCGTATGCCCT 4796
      |||||
Qy 106 tAlaAlaIaIaSerThrvAlaValIgluAspIleThrvIySasnIyGlnIuIyThrvSerG 126
      |||||
Db 4797 GCAGCTAAGACACT-----GAGCGTCAGCGTGTGAGAGCGCGAGATGAGCTGCA 4847
      |||||
Qy 126 uIleIle-----GluGlnIleProAlaSerGluIuSasnValIgluIuIyMetVa 142
      |||||
Db 4848 GCGCTCGGTGACAGCGCGTGCAGAGGACAGAAAGCGCGCGTGCAGAGAGAAC 4907
      |||||
Qy 142 IgluProIaIuIySerGlnAlaIaIaIyGlyPheIySvalIyPheIySphaValG 162
      |||||
Db 4908 CGAGGCTTGCAGAGCGAGTGCAGAGATGAGAGCCCAAGCAAGCGAGCGAGCGCA 4967
      |||||
Qy 162 yPheIySphThrvAlaIySvalSASP-----LysasnIyIySerAspThrvAlaI 179
      |||||
Db 4968 GCTGCGCTGCTGTGAGGACGAGAGCGAGAGCGAGCGCGAGAGAGCGCGCGCTGCA 5027
      |||||
Qy 179 nIeuIeuThrvAlaIySvalSASPgluIyGluIyAlaIuIy-----AlaSe 194
      |||||
Db 5028 GCGTCTGATGAACGAACTGCAGAGCGAGAGCGCGAGCGTGTGCTGCAAGCGCA 5087
      |||||
Qy 194 rValIyAlaIyAspHieGlnIuIyProSerValIuIyThrvAlaIyGly----- 210
      |||||
Db 5088 GGCAGAGAGGCTCCCAAGTGCAGGTAGCCTGTGAGACAGCGCAGCGTATGAGAAAGT 5147
      |||||
Qy 211 -GluSerAlaSerIySgluSerGluIeuIySgluSerThrvIySgluIyGlnIyThrvIle 230
      |||||
```



```

Oy 869 IHisThrValSerValAlaValIleAspGlyThrArgAlaValIleThrSerValGluGluArg 889
Db 7147 -----CTCAAGGCTCGCATTGAAGCTGAAACCG 7175
Oy 889 GSerProSerTrpIleSerAlaSerValThrGluPro---LeuGluHisThrAlaGlyI 908
Db 7176 GGCATCTATCTCTTCGTGCACAGACACACACAGCGCTTCCTCGAGAGAGAGCGGACGAA 7235
Oy 908 uAlaMetProProValGluGluValThrGluLysAspIleIleAlaGluGluThrProVa 928
Db 7236 GATGAAACAGGTGGCGACAGAGAAAGCTGCACGCTTGAGCGTGAAGCTGCCAGAGAGCAAG 7295
Oy 928 IleuThrGln-----ThrLeuProGluGluLysAspAlaHisAspMetVa 944
Db 7296 GCTGGGAGCTAGCCGAGAGAGACCTGGCCGACAGCGCGGCTCGCGAGAAAGATGCT 7355
Oy 944 IHisSerGluValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLe 964
Db 7356 GAAGAGAAAGATGCAGCGCGGTGCAGAAAGCCACAAAGCTCAAGCTGAGCTGAGCTGCT 7415
Oy 964 uArgThrGluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSerAl 984
Db 7416 GCAGCAGCAGAAAGAGAGCTG-----GCACAGAGCAGAGCGCGCGCTGCAGAGC 7463
Oy 984 aValSerGluLeuThrAspSerProAspThrThrGluGluAlaThrProValGluGluVa 1004
Db 7464 G-----GACAAAGAGCAAAATGCTCAGCACTGAGTGAAGGA 7499
Oy 1004 IGluSerGlyValLeuAspThrThrGluGluGluGluArgGlnThrGlnAlaIleLeuGlnAl 1024
Db 7500 GACACAGGCTTCCAGCGGAGACCTGTGAGGCTAGCGCGGCACTGAAATGAGCC 7559
Oy 1024 aValAlaAspLysValLys-----ValProAlaThrGlnThr-----ValGlnArgThrGl 1045
Db 7560 AAGAGCTGAACCGCTCAAGTTCGCGATGCGATGAGATGAGCCGAGCTCAGGCGCTGCAGA 7619
Oy 1031 -GluGluSerGln-----ValLeuAlaSerGluLysGluLysAspValMe 1071
Db 7630 GGAGAGTCCCGAGCGCTTCCGAGAGCGCTGAGAGATCGCGAAAGCTGCACCGCAC 7679
Oy 1035 -----ValProAlaThrGlnThr-----ValGlnArgThrGl 1045
Db 7680 TGAATCGCTACACAGAGAGAGATGACATTGGTCAGACTTCGAGATCCAGCGACAGAGA 7739
Oy 1045 ySer-----LysAlaLeuGluLysValGluGluValGluGluAspSerGl 1060
Db 7740 GAGTGACCAAGATGCCAGCGCTGTGAGGAGCGCATGCTGAGCTGAGACGAGAGA 7799
Oy 1060 u-----ValLeuAlaSerGluLysGluLysAspValMe 1071
Db 7800 GAAGCTCAAGCAGAGCGAGAGTACTGCAAGCTCAAGTCTGAGAGATGCAACATCTGTGA 7859
Oy 1071 tProLysGlyProValGlnGluAlaGlyAlaGluHisLeuAlaGlnGlySerGluThrGl 1091
Db 7860 GCAGAGCAGATATCTGCAGAGAGACACAGGCGCTGCAGAAAGCTTTCTCTGAGAGAGA 7919
Oy 1091 yGlnAlaThrProGluSerLeuGluValProGluValThrAlaAspValAspHisValAl 1111
Db 7920 CAGCTTGCGCAACCGGAAGCTTCATCGACGAGGAAGGCAAGCTGAGAGAGCTTTT 7979
Oy 1111 aThrCysGlnValIleLysLeuGlnGlnLeuMetGlu-----GlnAl 1125
Db 7980 CCAGAGCAGAGTGGCAAAAGCAAAACAGCTGCAGAGGAGGAGCAGCAGCGCAGAGAGAGA 8039
Oy 1125 aValAlaProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrProle 1145
Db 8040 GATGAGACGAGAAAGCAGAGAGCTGTGGCCAGCATGGAG-----GAGGCCGAGAG 8090
Oy 1145 uAlaAspSerAspThrAlaAspGlyThr-----GAGGCCGAGAG 8090
Db 8091 GGGGCAAGCTGAGGAGAGAGAGGTGTGAGCGCCAAAGCAAGAGAACTGACAGGCTGTGA 8150
Oy 1155 -----GlnGlnAspGluThrIleAspSerGlnAspSerLysAlaThrAlaAlaVa 1171

```

```

Db 8151 GCAGCAGCGGACGACGAGAGAAACTACTGCGCAGAGAGAGAACCAAGAGCTCGCGGAGCG 8210
Oy 1171 lArgIleSerGlnValThrGluGluGluAlaAlaThrAlaGlnLysGluGluProSerTh 1191
Db 8211 GCTCAGAGCGCTTGAGAGAAAGCAGCCAGCTGCTGGCCGCACTGTGAGAGATGCGCAC 8270
Oy 1191 t-----LeuProAsnAsnValProAlaGlnGluGluHisGl 1203
Db 8271 CTCCAGGCTGCTGCCAAGAAAGCATGCCCAT----- 8304
Oy 1203 yGluGluProGluArgAspValLeuGluProThrGlnGlnGluLeuThrAlaAlaAlaVa 1223
Db 8305 -----GCGCGCAGCGCATTTGATGCGCCCTCCATGGAG----- 8337
Oy 1223 lProValLeuAlaLysThrGluValGlyGlnGluGlyGluValAspThrPleuAspGlyI 1243
Db 8338 -----GCCGAGCCCGAGTACCTTTGAGAGGATTACGTGAGAAAGTGCAGCTCA 8387
Oy 1243 uLysValLysGlu-----GluGlnGluValPheValHisSerGlyProAsnSe 1259
Db 8388 GCAGCTCAGAGAACAGGCAATTCTGAGCATGAGAGAACTGCAGCGCTTGCACAGAGGTCA 8447
Oy 1259 rGlnLysAlaAlaAspValThrTyraAspSerGluVal-----Me 1272
Db 8448 CACCAAGGTGCTGAGGCTCAGCAGCGGAAAGATGTGCGCCACTGAAGGCGGCGAG 8507
Oy 1272 tGlyValAlaGlyCysGlnGluLysGluSerThrGluValGlnSer----- 1287
Db 8508 CAGCATGCGCAGATTTGCTCTGMAAGCCACCAATGAGAACTGAGTGTCTACAGAGCCT 8567
Oy 1288 -----LeuSerLeuGluGluGluGluMetGluTh 1297
Db 8568 ACAGCGCAGCTGTCAAGCCCTGGAACAGCCCTTATCTTACTTGAAGGCCAGAGCCTC 8627
Oy 1297 t-----AspValGluLysGluLysArgGluThrLysProGluGluLysSerGl 1313
Db 8628 GGGCTTCTGCTGACCTGTCGCCAGACCGGCGCTGACGCTCAATGAGCTGTGAAGA 8687
Oy 1313 u----- 1313
Db 8688 GGGTGTGTTGGGTCCCGAGCTGCACACAAAGCTGTGTCACTGAGCGTGCCTGATCG 8747
Oy 1314 -----GluGlyGluGlnGluThrAlaAlaProGluHisGlu----- 1325
Db 8748 CTACAAAGACCTTTACACAGAGAGAAAGAT--CTCTCTTCCAGGCGCATGAAGAGACC 8806
Oy 1326 -----GlyThrTyrgLysProValLeuThrLeuAsp----- 1336
Db 8807 TCATTGTGAGGACCATGTGCATCCGCTGTGGAAGCCCAAGATGCCACAGGTGGCATCA 8866
Oy 1337 -----MetProSerSerGluArgGlyLysAlaLeuGlySerLeu----- 1349
Db 8867 TTGACCGCTGTACACAGCAGCGGTTCCTGAGACGTGCGCTTCCAGCGTGGTACTTGG 8926
Oy 1350 ---GlyGlySerProSer-Leu-----ProAspGlnAspLysAlaGlyCysIleGluV 1366
Db 8927 ATGAGGAATGAACCGTGTGCTGCTGAGCCCAAGGATGACACCAAGGCGCTTTTGAAC 8986
Oy 1366 aGlnValGlnSerLeuAspThrThrValThrGlnThrAlaGluAlaValGluLysValI 1386
Db 8987 CCAACACTCAC-----GAGAACTCA 9007
Oy 1386 lGluThrValValIleSerGluThrGluGluSerProGluCysValGlyAlaHisLeuL 1406
Db 9008 CGTACCTGACAGTGTGAGAGCGGTGTGAGAGACCCGAG--ACAAGCTGCGGCTTCC 9064
Oy 1406 euProAlaGluLysSerSerAlaThrGlyGlyHis-Tyr---ThrLeuGlnHisAlaGlu 1424
Db 9065 TGCCACTCACAAGACAAAGCTGCAAGGCGTGTGAGCTGTGTACTGACACGAGAGCCC 9124
Oy 1425 AspThrValProLeuGlyProGluSerGlnAlaGluSerIleProIleIleValThrPro 1444

```

Db 9125 GTGACGCTCTCGAAGGCCACAGTGTCTGCACCATTCGCCAAGTTCCAGGGCAAGACCG 9184
 QY 1445 AlaProGluSerThrLeuHisProAspLeuGlnGlyGluLeuSerAlaSerGlnArgGlu 1464
 Db 9185 TGACCATCTGGGAGA-----TCATCAACTCAGAGTACT 9217
 QY 1465 ArgSerGluGluAspLysProAspAlaGlyProAspAlaAspGlyLysGluSerThr 1484
 Db 9218 TCACAGCGGAGCAGCAGCGGAGCTGCTCCGCGAGTTCCGCGC-----ATC 9270
 QY 1485 AlaileGluLysValLeuLysAlaGluProGluLeuLeuGluSerLysSerAsn 1504
 Db 9271 ACGTGGAGAGATCATCAGATTGTCATCAGCGTGGTAGGAAACAGCGCGAAGGCG 9330
 QY 1505 Lysile-ValLeuAsnValileGlnThrAlaValAspGlnPheAlaargThrGluThrAl 1524
 Db 9331 CAGCTCTGCTTTG----- 9343
 QY 1524 aProGluThrHisAlaTyraSerGlnThrGlnValProAlaCysArgLeuAspSerAr 1544
 Db 9344 -----AGGGCCTCGTGCCCTTGCTGTCTGCAGAGC-- 9376
 QY 1544 gGluProAsnArgCysTrpThrLysMetLysAspAlaLysMetLys 1559
 Db 9377 -----TGCTGGACA---GTGGAGTCATCATGTCATGAAG 9406

RESULT 10

US-09-954-456-1601
 ; Sequence 1601, Application US/09954456
 ; Patent No. US20020115057A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Young, Paul
 ; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
 ; SETS
 ; FILE REFERENCE: 689290-76
 ; CURRENT FILING DATE: 2001-09-18
 ; PRIOR APPLICATION NUMBER: US/60/233,617
 ; PRIOR FILING DATE: 2000-09-18
 ; PRIOR APPLICATION NUMBER: US/60/234,052
 ; PRIOR FILING DATE: 2000-09-20
 ; PRIOR APPLICATION NUMBER: US/60/234,923
 ; PRIOR FILING DATE: 2000-09-25
 ; PRIOR APPLICATION NUMBER: US/60/235,134
 ; PRIOR FILING DATE: 2000-09-25
 ; PRIOR APPLICATION NUMBER: US/60/235,637
 ; PRIOR FILING DATE: 2000-09-26
 ; PRIOR APPLICATION NUMBER: US/60/235,638
 ; PRIOR FILING DATE: 2000-09-26
 ; PRIOR APPLICATION NUMBER: US/60/235,711
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: US/60/235,720
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: US/60/235,840
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: US/60/235,863
 ; PRIOR FILING DATE: 2000-09-27
 ; NUMBER OF SEQ ID NOS: 2276
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 1601
 ; LENGTH: 14800
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-954-456-1601

Alignment Scores:

Pred. No.: 6,73e-20 Length: 14800
 Score: 418.00 Matches: 334
 Percent Similarity: 35.65% Conservative: 255
 Best Local Similarity: 20.22% Mismatches: 701
 Query Match: 5.18% Indels: 364
 DB: 10 Gaps: 61

US-09-902-432-4 (1-1596) x US-09-954-456-1601 (1-14800)
 QY 5 SerSerThrGluGlnArgSerProGlu-GlnProAlaGlySerAspThrProSerGluLe 24
 Db 4187 AGCAGCAGCGGGCAGAGGAGCGCGCGCTGCGCGAGGTGGAGGCGCGCTGAGAGAAGC 4246
 QY 24 uValLeuSerGlyHisGly-----ProAlaAlaGluAlaSerGlyAlaAlaGlyAspPr 42
 Db 4247 AGCGGAGCTGCGCGAGGCGCAGCGCCCAAGGCAAGGCGGAGCGGAGGCGAAGG 4306
 QY 42 oAlaAspAlaAspProAlaThrLysLeuProGlnLysAsnGlyGlnLeuSerSerValAs 62
 Db 4307 AGCTGAGCAGCGCCATGCA-----CGAGAGGTGTGTGCGCGGAGGAGCGG 4354
 QY 62 nGly-ValAlaGluGlnGlyAspValHisValGlnGluAsnGlnGluGlyGlnGlu 82
 Db 4355 CGGTGACGCGCAGCAGCAGCAAGCGCAGCATTCAGGAGGAGCTGCAGCAGCTCGCGCAGA 4414
 QY 82 luGluValValAspGluAspValGlyGlnArgGluSerGlu----- 95
 Db 4415 GCTCGGAGCGGAGATPCCAGGCCCAAGCGCCGCGCAGCAGAGCGCGCTGAGCGCAGCGGC 4474
 QY 96 --AspValArgGluLysAspArgValGluGluMetAlaAlaAsnSerThrAlaValGluA 115
 Db 4475 TGGCATCGAGGAGGAGATCCGGTGTGTGCGCTGCAGTTGGAGGCCACCGCAGCGCAGC 4534
 QY 115 spileThrLysAspGlyGlnGluGluThrSerGluile---lleGluGlnIleProAla 134
 Db 4535 GTGGCGGGGCTGAGGGGAGCTCCAGGCCTCGCTCACGGCGGAGGAGCTGAGGCAC 4594
 QY 134 erGluAsnAsnValGluGluMetValGlnProAlaGluSerGlnAlaAsnAspValGlyP 154
 Db 4595 AAAAGCAGCAGCGCAGGAGGAGCGCGCTTGGCGGAGCGAGGTGCAGGACGAGAGCC 4654
 QY 154 helysLysValPheLysPheValGlyPheLysPheThrValLysLysAsp-----L 171
 Db 4655 AGCGTAAGCGGCGAGCGGAGGTGGAGCTGCGCTCGCGCTGAAGCGCGCAGCAGCGCGG 4714
 QY 171 yeAsnGluLysSerAspThrValGlnLeuLeuThrValLysLysAspGluGlyGluGlyA 191
 Db 4715 CGCGCGAGAAGCAGCGGGCCCTCCAGGCCCTTGAGGAGCTGCGGCTGCAGCGGAGGAGG 4774
 QY 191 laGlu-----AlaSerValGlyAlaGlyAspHisGlnGluProSerValG 206
 Db 4775 CGAGCGCGCGCTCGCGCAGCGCGAGGTGCAGCGCGCGAGGTACAGGTGCGCTGG 4834
 QY 206 luThr-----AlaValGlyLysSerAlaSerLysGluSerGluLeuLysGln 222
 Db 4835 AGACGCGCAGCGCAGTGCAGAGCGCGAGCTGCAGAGCAAAACGCGCTCTCTCGCCGAGA 4894
 QY 222 erThrGluLysGlnGluGlyThrLeuLysGlnGluGln---SerSerThrGluileProL 241
 Db 4895 AGACGCGCAGCTGGAGCGCTCCCTGCAGGAGGAACACGTGCTGTGGCACACAGCTCGGG 4954
 QY 241 euGlnAlaGlu-----SerAspGlnAlaAlaGluGluAlaLysAspGluGlyG 258
 Db 4955 AGGAGCTGAGCGCGCGGCGCACAGCAGCGCGCGCGCGCGCGCGCGAGGAGCAG 5014
 QY 258 luGluLysGlnGluLysGluProThrLysSerProGluSerProSerProSerProValAsn 278
 Db 5015 AGCGGAGCTGGAGCGCTGCGCAGCTCAAGGCCCAAGCGCGCTACGCGCTGCGCGCTCAGG 5074
 QY 278 erGluThrThrSerSerPheLysLysPhePheThrHisGlyTrpAlaGlyTrpArgLysL 298
 Db 5075 CGGAGGAGGTGGCGCAGCAGAGAGCGCTG-----CGCG 5107
 QY 298 ysThrSerPheLysLysSerLysGluAspLeuGluThrAlaGluLysArgLysGluG 318
 Db 5108 AGGCCCGAGCTGAGAGCAGAGAGGAGGCGGAGCGCGCGCGCGCGCGCGCGCGCAGG 5167
 QY 318 lnGluAlaGluLysValAspGluGluGluLysGluLysThrGluProAlaSerGluGlu 338
 Db 5168 CGGAGGAGCAGCGCGTCCGCGCAGCGGAGTGGCTGAACAAGAG---CTGAGAGAGCAGC 5224

QY 338 InGUProAlaGUAspThrAspGlnAlaArgLeuSerAlaAspTyrGluValGlu 358
 Db 5225 GGCACCTGGCGAGAGCGACCGCGACAGCGCTGCGCCCGCGAGGATTGATCCGCGC 5284
 QY 358 euProLeuGluAspGlnValGlyAspLeuGluAlaSer---SerGluGluValCysAlaP 377
 Db 5285 TCGGGGCGCCAGACGAGCGAGCGAGCGACAGCGGCGACGCTGCTGGAGGAGAGCTGGGCC 5344
 QY 377 roLeu-----AlaThrGluValPheAspGlu----- 385
 Db 5345 GCGTCGACGCTGAGCGCGCTGCAGCGACCGAGAAACGCGAGGAGCTGAAGCGAGCTGG 5404
 QY 386 --LysMetGluAlaHisGlnGluVal-----LysGluValH 398
 Db 5405 CCAAGGTCGCGCGCGAGATGAGGTGCTGCTGCGCCAGCGAGCGAGCGCTGAAGGAGCT 5464
 QY 398 iValSerThrValGluValThrGlnGluGluGlnGlyGlyGlyGlyGluVal 418
 Db 5465 CGCGCTCCACCGAGAGAGTCCAGACGAGGCTG-----GAGCGCGAGGCGC 5512
 QY 418 lValValValGluGluThrGlyGlu-----SerLeuProProGlu--- 431
 Db 5513 GCGGCTTCGCGAGCTGCGCGCGAGAGCGCGCGCTGCGCTGCGCGCGCGAGAGGCCA 5572
 QY 432 -----LysLeuAlaGluProGlnGluValProGlnGluAlaGluProAlaGlu 448
 Db 5573 AGCGGAGCGCGAGCTGCGCGAGAGAGCGCGCGCGCGCGCGCGCGCGAG-----GCGGAGC 5629
 QY 448 lLeuMetLysSerArgGluMetCysValSerGlyGlyAspHisThrGlnLeuThrAspL 468
 Db 5630 GGGTGTTCGCGAGAGCGTGGCGCGCGCATCGC-----GAGGCCAGCGCGC 5674
 QY 468 euSerProGluGluValThrLeuProLysHisProGluGlyLysValSerGluValGlu 488
 Db 5675 TCAAGACGAGCGCGAGATCGCGCTCAAGAGAGAGCGCGAGAGCGCGCGCTGCGCGC 5734
 QY 488 etLeuSerSerGlnGlu-----ArgLysValGlnGlySerProLeuL 503
 Db 5735 GCGTGGCGAGAGCGAGCGCTTCAGCGCGCGCGCTGAGAGCGAGCGCGCGCGACACA 5794
 QY 503 yAlLysLeuPheSerSer---SerGlyLeuLysLysLeuSerGly-----LysL 518
 Db 5795 AGGCTGACATCGAGAGCGCTGCGCGCGCGCATGCGCAAGCATGCGACGCGAGCTGGAGC 5854
 QY 518 yGlnLysGlyLysArgGlyGly-----GlyGlyAspGluGluP 531
 Db 5855 GGCAGAGAGGGCT-GGTGAGAGACAGCTGAGAGCGCGCGCGAGGTGAGAGAGATC 5913
 QY 531 roGlyGluTyrGlnHisIleHisThrGluSerProGluSerAlaAspGluH-----L 549
 Db 5914 CTGGC-----GCTGAAGGCGAGCTTCGAGAGAGCGCGCGCTGCGCAAGCGCGAG 5961
 QY 549 yGlnGluSerSerAlaSerSerProGluGlu-----ThrCysLeuGlu 567
 Db 5962 CTGGAGCTGAGCTGGAGCGCATCGCACAGCGAGAGACAGCTGCGCGACGAAGAGAG 6021
 QY 568 LysGluProLeuGluAlaProGluAsp----- 576
 Db 6022 CAGGCGGAGCTGGAGCGCGAGGCGAGCGCGAGCTGCGCGCGAGAGAGAGCGCGCGC 6081
 QY 577 GlYlGluAlaGluGluGlyThrThrSerAsp-----GlyGluLys 589
 Db 6082 CGTGAAGCTGAGAGCGCGTGCAGAAAGAGCTGGCGCGAGAGAGAGCGCGAGAGAGCAG 6141
 QY 590 LysArgGluGlyLysThrProTrpAlaSerPheLysLysMetValThrProLysLysArg 609
 Db 6142 CGGAAGCGCGCGCTGAGAGAGTGCAGCGCGCTGAAGAGCCAACTGAGAGAGCGCGCGCGC 6201
 QY 610 ValArgArgProSerGluSerAspLysGluGluLeuGluLysValLysSerAlaThr 629
 Db 6202 CTGCGGAGCGAGCGAGAGCGAGCGCGCGCGAGCTGCGCGCGCGCGAGAGCGCGC 6261

QY 630 LeuSerSerThrAspSerThrValSerGluMetGlnAspGluValLysThrValGlu 649
 Db 6262 CAGAACCGCGT-----CAGCGAGAGAGAGAGCAACCGCTTCGCGTGG 6306
 QY 650 GluGlnLysProGluGluProLysArgArgValAspThrSerValSerTrpGluAlaLeu 669
 Db 6307 CAGCAGAGAGAGAGAGAGCTACAGCAGCGCTGCAGCAGAGAGAGAGC-----CTGCTG 6360
 QY 670 lLecysValGlySerSerLysLysArgAlaArgLysAlaSerSerSerAspAspGluGly 689
 Db 6361 GACCAGCTGCCCGGCGAGCGCGCGCGCGCGCGCGCTGAGAGAGCGAGAGCGC 6420
 QY 690 -----GlyProArgThrLeuGlyLysAspSerHisArgAlaGluAlaSerLys 706
 Db 6421 CGGTCGAGCGGAGCGGAGCGCGCGCGCGCGCGCGCGCGAGAGAGAGCGCGAGCGG 6480
 QY 707 AspLysGluAlaGlyThrAspAlaValProAlaSerThrGlnGlnGlnAspGlnAlaGln 726
 Db 6481 CTGAAGCAGTGC-----GCAAGAGAGAGAGAGAGAGCGCGCGCGCTCAG 6522
 QY 727 GlySerSerSerProGluProAlaGlySerProSerGluGlyGlyValSerThrTrp 746
 Db 6523 GCACAGCGCGCTGAGAG----- 6540
 QY 747 GluSerPheLysArgLeuValThrProArgLysLysSerLysLysLeuGluLys 766
 Db 6541 -----AAGCTGCGCAAGAGAGCGCGAGCGCGAGCGCG 6570
 QY 767 lAlaGluAspSerSerValGluGlnLeuSerThrGlnLysProSerArgGluGluSer 786
 Db 6571 GCGCGCGCGGAGAGCGGAGAGCGCGCGCGCGCGCGCGAGAGAGAGAGAGAGAGAGAG 6630
 QY 787 TrpValSerLysLysPheLysPheProGlyArgArgLysLysArgAlaAspGlyLysGln 806
 Db 6631 ATGAGAGAGCATGAATAATTGCGCGAGAGAGCTGCGCGAGAGAGCGCGAGAGAGAG 6690
 QY 807 GluGlnAlaThrValGluAspSerGlyProValGlnLysGlnAspAspProAsnVal 826
 Db 6691 GAGCTGACAACTGCGG-----CTGCACTGAGAGAGAGAGAGAGAGAGAGAGAGAG 6729
 QY 827 ProAlaValAlaProLeuSerGluTyrAsnAlaValGluArgGlu-----Lys 842
 Db 6730 -----CACCAAGAGAGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6771
 QY 843 MetGlu-----AlaGlnLysAsnThrGluLeuProGlnLeuGluValAla 858
 Db 6772 GCGAGAGCGAGAGCGCGAGCGCGAGCGCGAGCGAGAGAGAGAGAGAGAGAGAGAGAG 6831
 QY 859 TyrValSer---GluGlnLeuSerLysThrLeuValHisThrValSerValAlaValIle 877
 Db 6832 CGCGTGCAGATGAGAGAGAGCTGAGAG----- 6858
 QY 878 AspGlyThrArgAlaValThrSerValGluGluArgSerProSerTrpLysSerAlaSer 897
 Db 6859 -----CTCAAGGCGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6912
 QY 898 ValThrGluProLeuGlnHisThrAlaGlyLysAlaMetProProVal---GluGluVal 916
 Db 6913 AATACGAGCGCTTCCTGAG 6972
 QY 917 ThrGluLysAspLysLysAlaGluGluThrProValLeuThrGln-----Thr 932
 Db 6972 GCGCGCTGAGTGTGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7032
 QY 933 LeuProGluGlyLysAspAlaHisAspAspMetValThrSerGluValAspPheThrSer 952
 Db 7033 CTGGCAG 7092
 QY 953 GluAlaValThrAlaThrGluThrSerGluAlaLeuArgThrGluGluValThrGluAla 972
 Db 7093 GAGGCGACGCGAGCTGAAGGCTGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7152
 QY 973 SerGlyAlaGluGluThrThrAspMetValSerAlaValSerGlnLeuThrAspSerPro 992

QY 602 s-----MetValThrProLysLysArgValArgPr 613
 Db 3783 AAATGCTAAATAATGATGTTGATCAAGCTGTGCACAACTCAAAATCAAGCAATTCATAATAC 3842
 QY 613 oSerGluSerAspLysGluGlu----- 620
 Db 3843 AACTGGTGCTACAACTGAAGAAAAAATGACGAAAAAGATTAGTTTAAAGCTAAAGA 3902
 QY 621 -----GluLeuLysValLysSe 627
 Db 3903 AAAAGCGTATCAAGATATCTTAAATGCACAAACAACTAACGATGTTACGCAAAATTAAGA 3962
 QY 627 rAlaThrLeuSer-----SerThrAspSerThrValSerGluMet---Gl 641
 Db 3963 TCAAGCAGTGTGCTGATGTTCAAGGTATTTACTGCAGATACACAACTAAAGATGTGCGAA 4022
 QY 641 nAspGluValLysThrValGlyGluGluGlnLys----- 652
 Db 4023 AGATGAATTAGCAACAAAGCAAGGAAACAAAGAAAGCACTTATTGGCAAAACTGCAGATGC 4082
 QY 653 -----ProGluGluProLysArgArgValAspThr-SerValSerTrp----- 666
 Db 4083 GACTACTGAAGAAAGAAACAAACAAATCAACAGTAGATGCAGATTAACGCAAGGTAA 4142
 QY 667 -----GluAlaLeuIleCysValGlySerSerLysLysArgAlaAr 680
 Db 4143 TCAAAATATTGAAATATGCACAGTCAATCGATGATGTAAACACACTGCAAAAGATAATGCAAT 4202
 QY 680 GlyAla-----SerSerSerAspAspGluGlyGlyProArgThr----- 693
 Db 4203 TCAAGCAATTTGACCCAAATTCAGCATATCAACAGATGTTTAAACGAATGCAAGCGGAATT 4262
 QY 694 -----LeuGlyGlyAspSerHisArgAlaGl 702
 Db 4263 GCTAACTGAATGCAAATTAATACTGAATACTTAATAATGAGACTACTAATGA 4322
 QY 702 uGluAlaSerLysAsp-----LysGluAlaGlyThrAspAlaVa 715
 Db 4323 AGAAAAAGGTAACGATATTGGACCAGTTAGCAGCTGCATATGTAAGAGGTTTAAATAATAT 4382
 QY 715 lProAlaSerThr-----GlnGluGl 722
 Db 4383 TAATGACGCAACTACTACAGGTGATGTAAGTCTGCTTAAAGATACAGCAGTACAAAAAGT 4442
 QY 722 nAspGlnAlaGlnGlySerSerProGluProAlaGlySerProSerGluGlyGluGl 742
 Db 4443 TCACAACTTCATGCAATCTCTGTTAAGAGCCAGCAGGTAAACTGCTTTAGATCAAGC 4502
 QY 742 yValSerThrTrpGluSerPhelLysArgLeuValThrProArgLysLysSerLysSerLy 762
 Db 4503 TGCAGCTGAT--AAGAAAAACAAATAGAACAAACACCAAAATGCGTCACAAAGAAAT 4559
 QY 762 sLeuGluGluLysAlaGlu---AspSerSerValGluGlnLeuSerThrGluLeuGluPr 781
 Db 4560 TAACGATGCAAAAACAAAGCTTGATCTGAATTAATCAAGCGAAACAAATATTGACCA 4619
 QY 781 oSerArgGluGluSerTrpVal-----Se 789
 Db 4620 ATCTTCAACGATGAATATGTTGATATGTCAGTTAAAGAGAAAGCTTAAATTAATGC 4679
 QY 789 rIleLysLysPheIleProGlyArgArg----- 798
 Db 4680 AGTTAAACACTTTAGTCAGTACAAAAGAGATGCTTTAGCTAAGATTCAGCTGCATATAA 4739
 QY 798 ----- 798
 Db 4740 TGCTAAAGTAACCGACGCGGATAACTCTAACGCGATCGACTTCAAGTGAATTCAGAAAGC 4799
 QY 799 -----LysLysArgAlaAspGlyLysGlnGluGlnAlaThrVa 811
 Db 4800 GAAACAAAAACTTGCTGAATTGAACAAACACGCGGATCAAAATGTGCAATCAAGCTACTTC 4859

QY 811 lGluAspSerGlyProValGluIleAsnGluAspAspProAsnVal----- 826
 Db 4860 TAAAGATGATTCATTGAAGTTCAAATTCATTAATGACTTAGATATATTAACGATTACACAAT 4919
 QY 827 -----ProAlaValValProLeuSerGluTyAsnAlaValGluAr 840
 Db 4920 TCCAACAGGTAAAAAAGAAATCAGCTACACAGATTTATATGCTATGCGAGATCAGAAGAA 4979
 QY 840 gGluLysMetGluAlaGlnGlyAsnThrGluLeuProGlnLeuGluGlyAlaValTyVa 860
 Db 4980 AAATAATATTTTCAGCTGCACACTTAATGCAACCAAGATGAAGAACCAAGCAATTAAGCA 5039
 QY 860 lSerGluGluLeuSerLysThrLeuValHisThrValSer----- 873
 Db 5040 AGTTGACCAAAATGTTCAAACCTGCATTAAGAAAGCATTAATATGTTGGTGTGATAATCGTGA 5099
 QY 874 -----ValAlaValIleAspGlyThrArgAlaValThrSe 885
 Db 5100 CGTTGATGATGATTAACCAAGGTAAAGCAATTTGATGCTATTCAAGTAGATGCTAC 5159
 QY 885 rValGluGluArgSerProSerTrpIleSerAlaSerValThrGluProLeuGlu----- 903
 Db 5160 TGTAAACCTAAAGCGAACCAAGTGTATGCTAAAGCAGAGAAACCAAGAAATCTAT 5219
 QY 904 -----HisThrAlaGlyGluAlaMetProProValGluGl 915
 Db 5220 TGATCAAAGTGACCACTTAACCTGCTGAAGAAAAAAGCTGAAGCATTAGCAATGATTAACA 5279
 QY 915 uValThrGlu-----LysAspIleIleAlaGluGluThrProVal-----LeuThrGl 931
 Db 5280 AATTACAGATCAAGCTTAAACCAAGGTATTACTGATGCAACCAACACTGCTGAAGTTGAAA 5339
 QY 931 nThrLeuProGluGlyLysAlaHisAspMetValThrSerGluValAspPheTh 951
 Db 5340 AGCGAAAGCTCAAGACTTGAGCATTTGATTAACATT-----CAAATGACTCAAC 5390
 QY 951 rSerGluAlaValThrAlaThrSerGluAlaLeuArgThrGluGluValThrGl 971
 Db 5391 AGAAAAAAGCTATCGAAGAAATAGAACTGCACTAGACCAGATTGAAGCAGGTGT 5450
 QY 971 uAlaSerGlyAlaGluGluThrThrAsp----- 980
 Db 5451 AAATGTCAGCTGATGCTACCACTGAAGAAAAAAGCGTTTACGAATGCTTTAGAAGA 5510
 QY 981 ----MetValSerAlaValSerGlnLeuThrAspSerProAspThrThrGluAlaThr 999
 Db 5511 CATTTTCAAAAGCACTGAAGATATTTCTGATCAAACTACAAATGCGAGAAATCGCTAC 5570
 QY 999 rProValGlnGluValGluSerGlyValLeuAspThrGluGluGluGluArgGlnThrGl 1019
 Db 5571 T-----GTCAAAATAGTGCCTTGAACCACTTAAAGCAACGCTATTATATCC 5618
 QY 1019 nAlaIle-----LeuGlnAlaValAlaAspLysValLysGluGluSerGlnVa 1035
 Db 5619 AGTAGTTAGAAAAAATGCTTTGGAAGCAATTAGAGAAGTGGTTAAACAGCAAAATAGAAAT 5678
 QY 1035 l-----ProAlaThrGlnThrValGlnArgThrGly----- 1045
 Db 5679 AATTAAAAATGCGAGATGCGATCGCGGAAAGAAATTTGACGCTACAGATTTAGGTAG 5738
 QY 1046 -----SerLysAlaLeuGluLysValGluGluValGluGluAspSerGluValLeuAl 1063
 Db 5739 ATATTTTGACAGATTTGCTGATAAATTAGATAAAACACAAACAAATACAGAAGTT----- 5793
 QY 1063 aserGluLysGluLysAspValMetProLys-----GlyProValGlnGluAl 1079
 Db 5794 -GCTGAATTACAAAAATGTAACGATACCTGCAATTTGAAGCGATTGTCCTCAGAAATGATCC 5852
 QY 1079 aGlyAlaGluHisLeuAlaGlnGlySerGlu-----ThrGl 1091
 Db 5853 AAATGCTAATGATGATAATAGTGGTTTCAATTAATATGATGCAACAGCAAAATTCAAATGC 5912
 QY 1091 yGlnAlaThrProGluSerLeuGluValProGluValThr-----AlaAs 1106

```

Db 5913 TAAATGCAACACGAGAAATATCTGACACAACTTAATGTAATCTGACACAGATTAATGCCAA 5972
Qy 1106 pValasphIsValAlaThrCys-GlnValIleIysLeu----- 1118
Db 5973 TGCAGATACGCTCTTCAACAACGAAATATGATGATGCGGCAACAGTGAAACAAAC 6032
Qy 1119 -----GlnGlnLeuMetGlnGlnAlaValAlaProGlnUserSerG 1132
Db 6033 AGCAACATCAGCAAAAGTTGACGCAACTGATGATGCAAAATGATAAA-CCGCAAGCTAAATA 6091
Qy 1132 lUThrLeuThrAspSerGlnThrAsnGlySerThrProLeuAlaAspSerApThrAla 1152
Db 6092 ACAAT-----TCTAGTGAGATACATCAACAACAGCTCCGCAATGATGATGATGTA 6148
Qy 1152 spGlyThrGlnGlnAspGlnThrIleAsp-----SerGlnAspSerIValaThrAlaIav 1171
Db 6149 GCAAAACGAGATGAGATCAACTAATATACGCGCAAACTGATAACTGCAACAGAAACA- 6207
Qy 1171 aIArgGlnSerGlnValThrGlnGlnAlaAlaThrAlaGlnIysGlnGlnProSer 1191
Db 6208 -----GATTAATGCAACGCGCAAGCAAAAGCGCAAAATTAACAATAGTACAACTACAG 6259
Qy 1191 hrLeuProAsnAsnValProAlaGlnGlnGlnIleGlnIysGlnGlnProGlyArgAspVal 1211
Db 6260 CAACAAACGAGATGACCAACAGAA----- 6285
Qy 1211 euGlnProThrGlnGlnGlnIleuThrAlaAlaValaProValLeuAlaIysThrGln 1231
Db 6286 -----TCAACGCAACGCGCTCCAAACACTGCGAGACA----- 6318
Qy 1231 aIglyGlnIleGlnIysValaAspThrLeuAspGlyGlnIysValIysGln---GlnGln 1250
Db 6319 --GGTGCAAAATCATCTGCTGATGACCAAGATTAATCATCTGTAATGATTCACAAACAA 6376
Qy 1250 lValaPheValHisSerGly-----ProA 1258
Db 6377 AGCTCAAGTAAATTAATAGTGCATCTCAATCAACTAATGCAAGGTTGCAACACCA 6436
Qy 1258 snSerGlnIysAlaIaAspValThrTyraSpSerGlnValMetGlyValAlaGlyCysG 1278
Db 6437 AATCTGAAATTAAGCTAAGGCAAGAAAGATGTGAGATTCACAAATCAAGATATG 6496
Qy 1278 lngIulysGlnSerThrGlnValaGlnSerLeuSerLeuGlnGlnGlnMetGlnThra 1298
Db 6497 TTGAATCAACAACCTGAACATGCTTCAGACACATTAACAGAACCAAGATCTTCTTA 6556
Qy 1298 spValGlnIysGlnIysAspGlnThrIysProGlnIleValSerGlnIu-----G 1315
Db 6557 ATACATCAAAAGATTAAGAAAGAAAGTACGACAAAGTCAACAGATGCTGAGCAACACAAAT 6616
Qy 1315 lYglnGlnIleuThrAlaAlaProGlnHisGlnIleuThrIyrgly----- 1329
Db 6617 CTGATCAAAACGTTGCATCTTAACGAGACAGATTAATCTGAAGCAACGTTGATATGATG 6676
Qy 1330 -----LysProValLeuThrLeuAspMetProSerSerGlnIysGlyValIleuG 1347
Db 6677 TATCGAATTAACCA-----TCAACATCTAAACCGTCAAGAACCAAAAGATTAAGCG----- 6726
Qy 1347 lYSerLeuGlyGlySerProSerLeuProAspGlnAspIysAlaGlyCysIleGlnValG 1367
Db 6727 -----ACTTCAGCTAGAGATGTCAAAAAGGTGACATGCGCAACGCAACGCAAG 6769
Qy 1367 lValGlnSerLeuAspThrThrValThrGlnThrAlaGlnIleuValaValaGlnIysValIleG 1387
Db 6770 ATACAAAGATATATCAAGCATCAATGTCGACATGCAAGTGAATTAACATTAAGATACAG 6829
Qy 1387 lUThrValValIleSerGlnThrGlyGlnSerProGlnCysValGlyAlaHisLeuLeuP 1407
Db 6830 AAAAC-----GAAGGTGCGCAATGCACTCTC 6853
Qy 1407 roAlGlnIysSerSerAlaThrGlyGlyHisIleThrThrLeuGlnHisValaGlnAspThrV 1427

```

```

Db 6854 CAGCT-----ACAG 6862
Qy 1427 aProLeuGlyProGlnUser---GlnAlaGlnSerIleProIleIleValThrProAlaP 1446
Db 6863 TTTCAAAAGGAAGCAATGTGTCTAATCAAGATATGCTTAATGTAATCAAGACTTAAGAA 6922
Qy 1446 roGlnSerThrLeuHisProAsnLeuGlnGlnIysGlnIleSerAlaSerGlnArgIys 1466
Db 6923 ACAAAGCAAAATGCAAAATACGCCCAACAAAGAAAGTTAAT-----AAGCTTAAC 6973
Qy 1466 erGlnGlnIysAspIysProAspAlaGly 1475
Db 6974 AACAAAGTAAAACTTACCAAGTACTGCT 7002

RESULT 12
US-09-815-242-8291
; Sequence 8291, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zykkind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8291
; LENGTH: 7107
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1) ... (7107)
US-09-815-242-8291

Alignment Scores:
Pred. No.: 8.8e-20 Length: 7107
Score: 411.00 Matches: 325
Percent Similarity: 35.01% Conservative: 302
Best Local Similarity: 18.15% Mismatches: 671
Query Match: 5.09% Gaps: 495
DB: 10 Gaps: 72

US-09-902-432-4 (1-1596) x US-09-815-242-8291 (1-7107)

Qy 47 ProAlaThrIysLeuProGlnIysAsnGlyGlnLeuSerValAsnGlyValaIleGln 66
Db 2029 CCAAAGTTAAACCTCAAGCGCAAAACAGATATTAACCAAGTAACTGTAACAA 2088
Qy 67 GlnGlyAspValHisValGlnGlnIleuGlnGlnIysGlnGlnGlnIysGlnGlnGln 85

```


2089 CAA-----ATTAAAAAGTCAAAATGCATCATTACAAGATGAAAAAGATGTAGCA 2136
QY 86 AspGluaspValGlyGlnArgGluSerGluaspValArgGluLysasp----- 101
2137 AATGATAAAATTTGTAATTTGAAACAGCAATTTAAAGATATTGTCAGCAACAACA 2196
QY 102 ArgValGluGluMetAlaAAsenSerThrAlaValGluAspIleThrLysAspGlyGln 121
2197 AATGCACAAGTAGAAGCCATTAAACAAAGCGATCATGATATTATCAAACTGCACCT 2256
QY 122 -----GluGluThrSerGluIle-----Ile 128
2257 GCTACAACAGCTAAAGCAGCAGCTCTTGAAGAATTTGACGAAGTGTTCACGACAAAT 2316
QY 129 GluGlnIleProAlaSerGluAAsenValGluGluMetValGlnProAlaGluSerGln 148
2317 GATCAAGCACCTTTTAAATCTCTGACACAAATGAAGAAGTAGCGGAAGCTATTGAACGT 2376
QY 149 AlaAsn-----AspValGlyPheLysLysValPheLysPheValGlyPheLysPheThr 166
2377 ATTATGCAGCTAAAGTTTCTGTTTAAAGCAATTGAACGGACAAGCTGCACAAAGAT 2436
QY 167 ValLysLysAspLysAsnGluLysSerAspThrValGlnLeuLeuThr----- 182
2437 TTAGAAAGAGCTTAAACAGCAAGAAATCTCAAAATTTGAAAATATTACTGACTCTACTCAA 2496
QY 183 ValLysLysAspGluGlyGluGlyAlaGluAlaSerValGlyAlaGlyAspHisGlnGlu 202
2497 ACAAAATGATGCTATATGATGAATTAACAGACAGCAACAGCTAGAAAAAATCAAAAT 2556
QY 203 ProSerValGluThrAlaValGlyGluSerAlaSerLysGluSerGluLeuLysGlnSer 222
2557 GCTACAGTTTCAAATGCAATTAATGAAGAAGTAGCAGAGCTGATGACAGTAGAAGCA 2616
QY 223 ThrGluLys-----GlnGluGlyThrLeuLysGlnGluGlnSerSer 236
2617 GCTCAAAAGCAAGTTTACATGACATCCAAAGTTGTTAAATCAAAACAGAGAGTTGCTGAT 2676
QY 237 ThrGlu-----IleProLeuGlnAlaGluSerAspGln 247
2677 ACAAAATCAAAAGTATTAGATAAAATCAATGCAATTCAAACACAGCAAGCAAAAGTTAAACCT 2736
QY 248 AlaAlaGluGluAlaLysasp----- 255
2737 GCAGCTGATACGGAAGTAGAAGACGCATATAATACAGTAAACAAAGAAATTCAAAATAGC 2796
QY 256 -----GluGlyGluGluLysGln-----GluLysGluProThrLysSerPro 269
2797 AATGCTTCAACTACAGAAAGAAACAAAGCTGCATATACAGAAATAGACTATAAAGACAA 2856
QY 270 GluSerProSerProValAsnSerGluThrThrSerSerPheLysLysPhePheThr 289
2857 GAAGCAAGAAACAAATCTTGATGCTGCAAAATACAAACAGATGATGTAACACAGCTAAAGAC 2916
QY 290 HisGlyTrpAlaGlyTrpArgLysLysThrSerPheLysLysSerLysGluAsp-----Asp 308
2917 AATGGTATTCTGCTCAATTAATCAAGTACAGCGGCAACAACTAAGAAATCGGATGCTAAA 2976
QY 309 LeuGluThrAlaGluLysArgLysGlnGlnAla-----GluLysValAspGluGlu 326
2977 GCGGAATCTGCTCAAAAGCAAGTAGAAGCAAGTAAACTGCAATTCGAACCAATGAATGATCG 3036
QY 327 GluLysGluLysThrGluProAlaSerGluGluGlnGluProAla----- 341
3037 ACTACTGAAGAACAAACAGCTGCAAAAGATAAAGTTGATCAAGCAGTAGTTACTGCAAAC 3096
QY 342 GluAspThrAspGlnAlaArgLeuSerAlaAspTyrGluLysValGluLeuProLeuGlu 361
3097 GCTGATATAGATAATGCTGCGCAAAATACTGATGATGATAATGCAAAAAAATCTAATGAA 3156
QY 362 AspGlnValGlyAspLeuGluAlaSerSerGluGluLysCysAlaProLeuAlaThrGlu 381
3157 GCTACAAATCGCAGCCATTACACCTGATGCAAAATGTTTAA-----CCAAACAGCGAAACAA

QY 382 ValPheAspGluLysMetGluAlaHisGlnGluValValAlaGluValHisValSerThr 401
3211 GCAATTGCTGATAAAGTACAGCGCAAGAAACAGCAAT----- 3249
QY 402 ValGluLysThrGluGluGlnGlnGlyGlyGlyGlyGlyGlyGlyValValVal 421
3250 -----GATGCTAATTAACCGTGCACAACA 3273
QY 422 GluGlyThrGlyGluSerLeuProGluLysLeuAlaGluProGlnGluValProGln 441
3274 GAA-----GAAAAAGCAGCTGCGAAAAACAACAAAGTTCAAACT 3309
QY 442 GluAlaGluProAlaGluLeuMetLysSerArgLysMetCysValSerGlyGlyAsp 461
3310 GAAAAACAACAGCTGATACAGCAAT-----GATGGTGCA 3345
QY 462 HisThrGlnLeuThrAspLeuSerProGluGluLysThrLeuProLysHisProGluGly 481
3346 CATACAAAT-----GCAGAGTTGAAGCGCTAAAAATGCAGAA----- 3384
QY 482 IleValSerGluValGluMetLeuSerSerGlnGluArgIleLysValGlnGlnSerPr 501
3385 ---ATTGCTAAATTTGAAGCAATTCAGCCAGCAACAACAATAAGATTAATGCGAAAC-- 3439
QY 501 IleLysLysLeuPheSerSerSerGlyLeuLysLysLeuSerGlyLysGlnLysGln 521
3440 ---AAGCAATTTGCTACGA-----AAGCGAATGAACGTAAAAACAGCAATTCGCT 3483
QY 521 YLysArgGlyGlyGlyGlyAspGluGlu-----ProGlyGlyGlyGlnHisIle 537
3484 CAAACCGAAGACATTTACTGCTGAAGAAATTCGCGGCAATTCGCAATGTAGATAATGC 3542
QY 537 eHisThrGlu-----SerProGluSerAlaAspGluGln 548
3543 TGTGACACAGCAAAATAACAACATTCAGCTGCTAATAGTCAAAATGATGTAGACCAAGC 3602
QY 548 nLys-----GlyGluSerSerAlaSer-----SerProGluGluProGluGluThrTh 564
3603 GAAAAACAACCTGGTGAAGCTAGTATTGATCAAGTAAACACCAACAGTAAATAAAGCAAC 3662
QY 564 rCysLeu-----GluLysGlyProLeuGluAlaProGlnAspGlyGluAlaGluGlu 582
3663 AGCAGTACAGACGCGGCAAAATAATATTACAGCTGCACTGATGATGATGTTGTTAGATAC 3722
QY 582 YThrThrSerAspGlyGluLysLysArgGlyIleThrProTrpAlaSerPheLysIle 602
3723 AGCGAAAGATGAGGTAAAAATTCGATTCAAGAGTACACACCAGCAACAGCGGTAAATC 3782
QY 602 s-----MetValThrProLysLysArgValArgArgPr 613
3783 AAATGCTAAAAATGATGTTGATCAAGCTGTGCAACTGCAAAATCAAGCAATTTGATAATAC 3842
QY 613 oSerGluSerAspLysGluGlu----- 620
3843 AACTGGTGTACAACTGAAGAAAAAATAATGCACAAAAGATTTAGTTTTAAAAAGCTAAAGA 3902
QY 621 -----GluLeuGluLysValLys 627
3903 AAAAGCGTATCAAGATATCTTAAATGCACAAACAACTACGATGTTACGCAAAATTAAGA 3962
QY 627 rAlaThrLeuSer-----SerThrAspSerThrValSerGluMet---Gln 641
3963 TCAAGCAGTTGCTGATGTTCAAGGTATTACTGCAGATACAACAATTAAGATGTTGCGAA 4022
QY 641 nAspGluValLysThrValGlyGluGluGlnLys----- 652
4023 AGATGAATTTAGCAACAAAAAGCAAGGGAACAAAAAGCACTTTATTCACAACATGCAGATGC 4082
QY 653 -----ProGluGluProLysArgArgValAspThrSerValSerTrp----- 666
4083 GACTACTGAAGAAAAAGAAACAAACAAATCAACAGTAGATGCAAGATTTAAACCAAGGTAA 4142

QY 667 -----Glu1aLeu1IecyVal1GlySerSerlyLysArg1aAr 680
 Db 4143 TCATAATATTGAAATGACAGTCATGATGTAACACTGCAAAAGATTAATGCAAT 4202
 QY 680 GlyS1a-----SerSerSeraspGluGly1ProArgThr----- 693
 Db 4203 TCACGCAATGACCCCAATTCACAGCATCAACAGTGTAAACAGAACGAAAGCGCAATT 4262
 QY 694 -----LeuGlyGlyAspSerHisArg1aG1 702
 Db 4263 GCTAAGTGAATGCAAAATATAAATACTGAAATTAATTAATTAATGAGACTACTAATA 4322
 QY 702 uGlu1aSerlyAsp-----LysGlu1aGlyThrAsp1aVa 715
 Db 4323 AGAAAAAGGTAAACGATATTGGACCACTAGAGTGCATGTGAAGAAAGTTTAATATATAT 4382
 QY 715 lPro1aSerThr-----GlnGluG1 722
 Db 4383 TAATGCAAGCACTACTACAGTGATGTAATCTGCTAAAGATACAGCACTACAAAAGT 4442
 QY 722 nAspGlu1aGlnGlySerSerSerProGluPro1aGlySerProSerGluGlyGluG1 742
 Db 4443 TCAACAACTTCATGCAATCCCTGTAAAGAGCCAGCAGGTAAACTGCTTAGATCAAGC 4502
 QY 742 yValSerThrTPGluSerPheLysArgLeuValThrProArgLysLysSerlySerly 762
 Db 4503 TGCAGCTGAT---AAGAAAAACAAATAGAACAAACCAACCAATGCTCAACAAACAAAT 4559
 QY 762 sLeuGluGlyLysValGlu---AspSerVal1GluGlnLeuSerThrGlu1leGluPr 781
 Db 4560 TTAACGATGCAAAACAAAGAGTTGATCTGATTAATTAATCAAGCAAAACAAATTTGACCA 4619
 QY 781 oSerArgGluGluSerTyrVal-----Se 789
 Db 4620 ATCTCAACAGATGATGATGTTGATATGCACTTAAGAAAGAAAGCTAAATTAATGC 4679
 QY 789 r1leLysLysPhe1leProGlyArg----- 798
 Db 4680 AGTTAAACATTTAGTAGATGACAAAGAAAGTCTTAGCTAAGATTGAAGCTGCATATTA 4739
 QY 798 ----- 798
 Db 4740 TGCATAAGTACCGAAGCGGATTAACCTTAACGATGACTTCAAGTCAAAATTCGACAAGC 4799
 QY 799 -----LysLysArg1aAspGlyLysGlnGluGln1aThrVa 811
 Db 4800 GAAACAAAACTGTGATTAATGAAACAAAGCGGATCAAAATGTCATCAAGCTACTTC 4859
 QY 811 lGluAspSerGlyProValGlu1leAsnGluAspAspProAsnVal----- 826
 Db 4860 TAAAGATGACATTGAAGTTCAAAATTCATATGACTTAGATTAATTAACGATTACCAAT 4919
 QY 827 -----Pro1aVal1ProLeuSerGluTyrAsn1aVal1GluAr 840
 Db 4920 TCCACACGCTAAAAAGATCACTACACAGATTTATATGCTATGACATGAGAAAGAA 4979
 QY 840 gGluLysMetGlu1aGlnGlyAsnThrGluLeuProGlnLeuLeuGly1aVal1YrVa 860
 Db 4980 AATATATATTTCAGTGCATATGCAACACAAAGATGAAACACAAAGCAATTAAAGCA 5039
 QY 860 lSerGluLysSerlyThrLeuValHisThrValSer----- 873
 Db 5040 AGTTGACCAAAATGTTCAAACTGCATTAGAAACATTAATTAATGCTGTGATATGCTGA 5099
 QY 874 -----Val1aVal1leAspGlyThrArg1aVal1Thrse 885
 Db 5100 CGTTGATGATGATTAAACACAGTAAAGCAATTTGATGCTAATTCATTAAGATGCTAC 5159
 QY 885 rValGluGluArgSerProSerThrPilesSer1aSerValThrGluProLeuGlu----- 903
 Db 5160 TGTATAACCTAAAGCGAAGCAAGTGAATGCTAAAGCAGAGGAAACGAAABATCTAT 5219
 QY 904 -----HisThr1aGlyGlu1aMetProProVal1GluG1 915

Db 5220 TGATCAAGTACCAAGTTAACTGCTGAAGAAAAAAGTAAAGCACTTAAGATGATTAAACA 5279
 QY 915 uValThrGlu-----LysAsp1le1le1aGluGluThrProVal-----LeuThrG1 931
 Db 5280 AATTACAGATCAAGCTAAACAAGGATATTACGTAGCAACAACACTGCTGAAGTTGAAGA 5339
 QY 931 nThrLeuProGluGlyLysAsp1aHisAspAspMetValThrSerGluValAspPheTh 951
 Db 5340 AGCGAAAGCTCAAGGACTTGAGCACTTTGATTAACATTT-----CAATTTGACTCAAC 5390
 QY 951 rSerGlu1aVal1Thr1aThrGluThrSerGlu1aLeuArgThrGluGluGluValThrG1 971
 Db 5391 AGAAACAAAAAGCTATCGAAGATTTAGAACTGCATTAAGCAAGATTAAGCAGGTGT 5450
 QY 971 u1aSerGly1aGluGluThrThrAsp----- 980
 Db 5451 AAATGTCAGCGTATGCTCAACTGAAGAAAGAAAGAGCGTTTACGAATGCTTTAGAGA 5510
 QY 981 ---MetValSer1aVal1SerGlnLeuThrAspSerProAspThrThrGluGlu1a1aTh 999
 Db 5511 CATTTATCAAAAAGCACTGAAAGATTTTGTGATCAACCTAACATGCAAGAAATCGTAC 5570
 QY 999 rProValGlnGluValGluSerGlyValLeuAspThrGluGluGluGluArgGlnThrG1 1019
 Db 5571 T-----GTCAAAAATAGTGGCGCTTGAACAACTTAAAGCAACGATTAATCC 5618
 QY 1019 nAla1le-----LeuGln1aVal1aAspLysVal1LysGluGluSerGlnVa 1035
 Db 5619 AGTATTTAAGAAAAATGCTTTGAAGCAATTAGAGAGTGTAAACAGCAAAATGAAT 5678
 QY 1035 l-----Pro1aThrGlnThrVal1GlnArgThrGly----- 1045
 Db 5679 AATTAAAAATGCAATGCAATGATCGCGCAAAATTCACGCTACAGATTAGAGTAG 5738
 QY 1046 -----SerLys1aLeuGluLysVal1GluGluVal1GluGluAspSerGluVal1Leu1 1063
 Db 5739 AATTTTGACAGATTTGCTGATTAATTTGATTAACACAAACAAATACAGAGTT----- 5793
 QY 1063 aSerGluLysGluLysAspValMetProLys-----GlyProVal1GlnGlu1a 1079
 Db 5794 -GCTGAATTACAAATATGTAACGATACCTGCAATGCAATGCAATGCGCTCAGATGATCC 5852
 QY 1079 aGly1aGlu1aLeu1aGlnGlySerGlu-----ThrG1 1091
 Db 5853 AATCTATATGATTAATTAATGATGTTGATGATTAATGATGATGATGATGATGATGATG 5912
 QY 1091 yGln1aThrProGluSerLeuGluVal1ProGluVal1Thr-----Alas 1106
 Db 5913 TTAATGCAACCAAGAAATATCTGCAACCACTTAATGTTAACTGATCAACAGATATGCCAA 5972
 QY 1106 pValAspHisVal1aThrCys-GlnVal1leLysLeu----- 1118
 Db 5973 TGCAGATACGCTTCAACACAGCAAAATATCAAGATGATGCGGCAACAGTGAACAAAC 6032
 QY 1119 -----GlnGlnLeuMetGluGln1aVal1aProGluSerSerG 1132
 Db 6033 AGCAACATGACGCAACAGTTGACGCAACTGATGATGATGATGATGATGATGATGATG 6091
 QY 1132 luthrLeuThrAspSerGluThrAsnGlySerThrProLeu1aLaspSerSerThr1a1a 1152
 Db 6092 ACAAT---TCTGTCAGATGATCAACAAACAGTCCGACAAATGATATGATGATGATG 6148
 QY 1152 spGlyThrGlnGlnAspGluThr1leAsp---SerGlnAspSerly1aThr1a1a1a1a 1171
 Db 6149 GCAAACCAAGATTAAGATCAATTAATTAACGCAACAATGATTAACCTGCAACAGAAAGCA 6207
 QY 1171 aLArgGlnSerGlnVal1ThrGluGluGlu1a1a1a1a1a1aGlnLysGluGluProSerT 1191
 Db 6208 -----GATATGCAACGCGCAGGAAAGCGCAACAAATTAACATATGATCAACTACAG 6259
 QY 1191 hLeuProAsnAsnVal1Pro1aGlnGluGlu1a1a1a1a1a1a1a1a1a1a1a1a1a1a 1211

```
Db 6260 CAACAAAGAGATGACACACAGAA----- 6285
Qy 1211 euGluProThrGlnGlnGluLeuThrAlaAlaValProValLeuAlaLysThrGluV 1231
Db 6286 -----TCAACAGCAACGGCTCCAACAACCTGCGAGCACA--- 6318
Qy 1231 alGlyGlnGluGluValAspTrpLeuAspGlyGluLysValLysGlu---GluGlnG 1250
Db 6319 --GGTGAGAAATCATCTGCTGATAGCAAGATATGATCTGTAATGATTCAAAACAAA 6376
Qy 1250 luValPheValHisSerGly-----ProA 1258
Db 6377 ATGCTGAAGTAATAATAGTCTCAATCTCAATCACTAATGCGAAGTTGCGACAAACAA 6436
Qy 1258 snSerGlnLysAlaAlaAspValThrTyrrAspSerGluValMetGlyValAlaGlyCysG 1278
Db 6437 AATCTGAATAATAAGCTTAAGGCAGAAAAAGATGTCAGATTTCAACAAATCAAAAGTATGG 6496
Qy 1278 lnGluLysGluSerThrGluValGlnSerLeuSerLeuGluGluGluMetGluThrA 1298
Db 6497 TTGAATCAACAACCTGAACATTTGCTTTCAGCAGACATACAGAACCAAAAGTATCTTCTA 6556
Qy 1298 spValGluLysGluLysArgGluThrLysProGluGlnValSerGluGlu-----G 1315
Db 6557 ATACATCAAAAGATAAAGAAAGTACGACAAGTCAACAGATGCTGAGCAACACAAATT 6616
Qy 1315 lyGluGlnGluThrAlaAlaProGluHisGluGlyThrTyrrGly----- 1329
Db 6617 CTGATACAAAGCTTGCTATTAACGACAGATAAATCTGAAGGCAACGTTGATACTGATG 6676
Qy 1330 -----LysProValLeuThrLeuAspMetProSerSerGluArgGlyLysAlaLeuG 1347
Db 6677 TATCGAATAAACCA-----TCACATCTAAACCGTCAGAACGCAAAAGATAAAGCG--- 6726
Qy 1347 lySerLeuGlyGlySerProSerLeuProAspGlnAspLysAlaGlyCyslleGluValG 1367
Db 6727 -----ACTTCGACTGAGGATAGTCAAAAGCTGACATGCGAACACAGCAG 6769
Qy 1367 lnValGlnSerLeuAspThrThrValThrGlnThrAlaGluAlaValGluLysValIleG 1387
Db 6770 ATACAAAAGATAATCAAGCATCAATGGTCAACTGCGAGATGTAACAATAAAGCTACGC 6829
Qy 1387 luThrValValIleSerGluThrGlyGluSerProGluCysValGlyAlaHisLeuLeup 1407
Db 6830 AAAAC-----GACGGTGCCTCAATGCATCTC 6853
Qy 1407 roAlaGluLysSerSerAlaThrGlyGlyHisTrpThrLeuGlnHisAlaGluAspThrV 1427
Db 6854 CAGCT-----ACAG 6862
Qy 1427 alProLeuGlyProGluSer---GlnAlaGluSerIleProIleIleValThrProAlaP 1446
Db 6863 TTTCAAAAGGAAGCAATAGTCTAATCAAGATATGCTTAATGTAACCTAAGCTAAAGAAA 6922
Qy 1446 roGluSerThrLeuHisProAspLeuGlnGlyGluIleSerAlaSerGluArgGluArgs 1466
Db 6923 ACRAAGCAAAATGCAAAATACGCCCAACAGCAAAAGTAAAT-----AAGCTTAAAC 6973
Qy 1466 erGluGluGluAspLysProAspAlaGly 1475
Db 6974 AACAGCTAAACCTTTACCAGATACTGGT 7002

RESULT 13
US-09-960-253-145
; Sequence 145, Application US/09960253
; Patent No. US20020123619A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Mohamath, Raedoh
; APPLICANT: Lodes, Michael J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.556
```

```
; CURRENT APPLICATION NUMBER: US/09/960,253
; CURRENT FILING DATE: 2001-09-20
; NUMBER OF SEQ ID NOS: 187
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 145
; LENGTH: 10300
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-960-253-145

Alignment Scores:
Pred. No.: 2,21e-19 Length: 10300
Score: 408.00 Matches: 315
Percent Similarity: 38.11% Conservative: 305
Best Local Similarity: 19.36% Mismatches: 670
Query Match: 5.05% Indels: 337
DB: 10 Gaps: 60

US-09-902-432-4 (1-1596) x US-09-960-253-145 (1-10300)
Qy 26 LeuSerGly-----HisGlyProAlaAlaGluAlaSerGlyAlaAla----- 39
Db 2376 CTCTCAGGTGAAGGAACCTTAGCATGCTTAACAGAAATTTAGGGCTCAGGTAAGCAACTGGA 2435
Qy 40 GlyAspPro-AlaAspAlaAspProAlaThrLysLeuProGlnLysAsnGlyGlnLeuSe 59
Db 2436 AATGAACCTTGCAGAGCAGAAAGCAAGACATTTGATTATGAAGC---CAAACATCG 2492
Qy 59 rSerValAsnGlyValAlaGluGlnGlyAspValHisValGlnGluGluAsnGlnGluG 79
Db 2493 CCATGACACCTGCTCCTCTGAACAG-----ATCCATAGTCTCAGCATAGAACCCAAATC 2546
Qy 79 yGlnGluGluGluValValAspGluAspValGlyGlnArgGluSerGluAspValArgG 99
Db 2547 TAAAGATGTGAATAAT-----GAAGTTTACAGAACTGAATGATGATGTCAG-- 2595
Qy 99 uLysAspArgValGluGluMetAlaAlaAsnSerThrAlaValGluAspIleThrLysAs 119
Db 2596 -----CTTCAGTTTTCTCAGCAGACAGTACCTCTAAGAAAGCCCTGCAAAAGCCA 2642
Qy 119 pGlyGlnGluGluThrSerGluIleIle-----GluGlnIleProAlaSerGluAs 136
Db 2643 GCTGCAAAATAGAAAGTGAAGTCTTGAAGGGGCGAGACGCTGTGAAGGCATATCTCAAG 2702
Qy 136 nAsnValGluGluMetValGlnProAlaGluSerGlnAlaAsnAspValGlyPheLysLy 156
Db 2703 TAAAGTGAAGAACTGCTCCAGGCTCTTTACAGAGAGAACTTGAATA----- 2751
Qy 156 sValPheLysPheValGlyPheLysPheThrValLysLysAspLysAsnGluLysSerAs 176
Db 2752 -----ACAAAAATCGAT----- 2763
Qy 176 pThrValGlnLeuLeuThrValLysLysAspGluGlyGluGlyAlaGluAlaSerValG 196
Db 2764 -----CAGCTCTTACTAGAAAAAGAGAGATGTGGAACCCCTCCACAAACCATCGA 2816
Qy 196 yAlaGlyAspHisGlnGluProSerValGluThrAlaValGlyGluSerAlaSerLysG 216
Db 2817 GGAGAGAGATCAACAAGTCACAGAAATCAGCTTTAGTATGACTGAGAAATGTTGCTCAGCT 2876
Qy 216 uSerGluLeuLysGlnSerThrGluLysGlnGluGlyThrLeuLysGlnGluGlnSerSe 236
Db 2877 TAATGAAGAGAAAGTTTTCTCTTGGGGTTGAAATTAAGACTCTTTAAAGAA----- 2925
Qy 236 rThrGluIleProLeuGluAlaGluSerAspGlnAlaAlaGluGluAlaLysAspG 256
Db 2926 -----CAGCTAAATTTATATCCAGAGCTGAGGAAGCAAAAAAGAGCAGCGTGAAGAAGA 2981
Qy 256 uGlyGluGluLysGlnGluProThrLysSerProGluSerProGluSerProProVa 276
Db 2982 TAATGAAGTTTCTTCTGGCCCTTAACAAATTTATGATGAGATGAGCCAGCAGCAAAAT 3041
Qy 276 lAsnSerGluThrThrSerSerPheLysLysPhePheThrHisGlyTrpAlaGlyTrpAr 296
```

```

Db 3042 AAGTAAAGGAAGAA-----CTTCAGCATGAATTGACCTTCTGAA 3080
Qy 296 gLySerThrSerPheLysSerLys-----GluAspAs 308
Db 3081 GAAAGAAATGAGCAGAGAAAGAAAGCTCCAGCAGCTTTATTACAGAAAGAGCT 3140
Qy 308 pLeuGluThrAlaGluLysArgLysGluGluGluAlaGluLysValAspGluGluLys 328
Db 3141 TCTGCAAAAGCTCAGTAGATTGGAGAAAGAAATTAGCCAACTTGAAAGAAATCTTAAGAA 3200
Qy 328 sGluLysThrGluProAlaSerGluGluGluGluProAlaGluAspThrAspGluAlaAr 348
Db 3201 AAAAAATCCACTCAGTGAAGCTGAGAGGGAGAGTGAAGAAAGATTAAGAAAGAAACAAAG 3260
Qy 348 gLeuSer-----AlaAspTyrGluLysValGluLeuProLeuGluAspGlu 363
Db 3261 ATAAGTCAAGAAATGCTGACTTCTTAAGTCCAGAAATTAAGAAATTTATTAAACAGAC 3320
Qy 363 nValGluAspLeuGluAla-----SerSerGluGluLysCysAl 376
Db 3321 AATATCTGAGAAAGAAAGTGAAGCTACAGCATATTAAGAAAGATTTGAGAAAGAAAG----- 3375
Qy 376 aProLeuAlaThrGluValPheAspGluLysMetGluAlaHisGluGluValAlaGlu 396
Db 3376 ---CTGGCAGCTGAA-----GAGCAATTCAGGCTCTGGTCAACAG 3413
Qy 396 uValHisValSerThrValAlaGluSerThrGluGluGluGluGluGluGluGluAlaGlu 416
Db 3414 GATGATCAGACCTTGCAAGATTAACAAACCA----- 3447
Qy 416 uGluGluValAlaGluGluGluGluGluSerLeuProProGluLysLeuAlaGluPr 436
Db 3448 -----ATGATTTGCTCCCAAGCAGAAATCGTGAAGAA 3479
Qy 436 oGluGluValProGluGluAlaGluProAlaGluGluLeuMetLysSerArgGluMetCys 456
Db 3480 CCAGCAATTTATCCAGAG-----TTATATCAAGATTAACAGCA 3518
Qy 456 sValSerGluGluAspHisThrGluLeuThrAsp-----LeuSerProGluGlu 472
Db 3519 TCCAGTGTATGGGAGCTCCGTAGCACTTTGAAGAAACAGTGTATAGTCCACTTG 3578
Qy 472 uLysThrLeuProLysHisProGluGluGluGluValSerGluValGluMetLeuSerSerGlu 492
Db 3579 TACAGGTAGTAGGACACTGGAACCAAGAACTAAGAAAGAAAGATCTGCGCTTGAA 3638
Qy 492 nGluArgGluLysValGluGluSerProLeuLysLysLeuPheSerSerSerGluLysLeu 512
Db 3639 AGAAAAAGGAGCAACTTCAAAAGAAAGCTACAGAGAGCTTAACTCCCGCAAGGCAATTC 3698
Qy 512 sLysLeuSerGluLysLysGluGluGluLysArgGluGluGluGluAspGluGluProGlu 532
Db 3699 TAAAAAGGACAGAGAAAGAAAGAAAGACATCTCAG-----GAGAGCTTAA 3743
Qy 532 yGluLysThrHisLysHisThrGluSerProGluSerAlaAspGluGluLysGluLysLeu 552
Db 3744 GCACAGAAAGATGACTATATGCTTGCAAGAACATTTGATGACAAAGCAAGGAGAA 3803
Qy 552 r-----SerAl 554
Db 3804 TGAGAAATATGAGACCAAGCTAAGCAACTCCAGATTCAAGTAAGGAAATCCATAGACG 3863
Qy 554 aSerSerProGluGluProGluGluThrThrCysLeuGluLysGluProLeuGluAlaPr 574
Db 3864 AAAAGCTCCAGACAGACAGCAGAGAAATCTGTCTTCCATCCAGCTTTAGAAAGAAC 3923
Qy 574 oGluAspGluGluAlaGluGluGluGluLysThr----- 584
Db 3924 TTTATTCAAGGACAGACAGATCAGACTCAACCTGTTTAGAGTCCAACTTGCGCC 3983
Qy 585 -----SerAspGluLysLysArgGluGluGluLysThrProThrAlaSer 599

```

```

Db 3984 AGACTGGCCTTCTCATTTCTGAAGATGCAAGTCTCTGACAGGCGGAATCTGTGCCCA 4043
Qy 599 rPheLysLysMetValThrProLysLysArgValArgArgProSerGluSerAspLysGlu 619
Db 4044 GATTAAAGCCCAAGCTG-----AAAGA 4064
Qy 619 uGluGluLeuGluLysValLys-----SerAlaThrLeuSerSerThrAspSerThrValSe 638
Db 4065 AATTAAGGCTGAGAAAGATTAAGATTAAGAAAGTTAGTTACTCAACACAGTGAAGCTTAC 4124
Qy 638 rGluMetGluAspGluValLysThrValGluGluGluGluLysProGluGluProLysArg 658
Db 4125 TAAAAATCAGAGAGATTTTCACTTCAACAGCAGATTAATAACAGGTTTAGAAAT 4184
Qy 658 gArgValAspThrSerValSerTyrGluAlaLeuLysCysValGluSerSerLysLysAr 678
Db 4185 CGAGAGCTTAAGACAGATATCCATGAAGCTGAAGTCCATGCCAGACCTCGACAGCA 4244
Qy 678 gAlaArgLysAlaSerSerSerSerAspAspGluGluGluProArgThrLeuGluGluAspSe 698
Db 4245 ATTG-----GAAAGCAGCCCACTACAAATGTGCTGCCTAGA 4280
Qy 698 rHisArgAlaGluGluAlaSerLysAspLysGluAlaGluLysThrAspAlaValProAlaSe 718
Db 4281 ACATCTAAGAGATTTGCAACCTAAACCTGATGACTG-----CAAAATCATTAAG 4331
Qy 718 rThrGluGluGluAspGluAlaGluGluSer-----SerSerProGluProAlaGlu 735
Db 4332 CAAAAGAGAGAGAGAGCTTACTCTTTCGACCACTTAGTGAGAAAGAGAGCTCT 4391
Qy 735 ySerProSerGluGluGluGluGluValSerThrTyrGluSerPheLysArgLeuValThrPr 755
Db 4392 CACTAAATTAACAGACAGATTAAGAACAAAGAAATTAATTAAGCTGTCATAGC-- 4449
Qy 755 cArgLysLysSerLysSerLysLeuGluGluLysAlaGluAspSerSerValGluGluLe 775
Db 4450 -----CAGCTAGAAATTCAGCCAAAGAGATATGAGAGATTAAGCACT 4496
Qy 775 uSerThrGluLe-----GluProSerArgGluGluSerTyrPa 788
Db 4497 ACAGGTGGAACCTTGTGAATGAAGCAAAACCAAGAGATGTGGAGAGAAAGT----- 4551
Qy 788 lSerLeuLysPheLeuProGluArgArgLysArgAlaAspGluGluGluGluGlu 808
Db 4552 -----AGAGCAAGCAACAAATTAACAAAGAAACCTGCACAGC 4586
Qy 808 nAlaThrValGluAspSerGluProValGluLysHisGluAspAspProAlaProAl 828
Db 4587 TGCCCTTATTTCCGAAAA-----GAAAGCACTAAAGAAAGAAAGAGTCTCAGAG 4637
Qy 828 aValValProLeuSerGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 840
Db 4638 GGAATTTGCTTTGGCCAGAA-----GGTACCATTAAGAGTCTCAACCAAGTCTCGCAGATG 4694
Qy 841 -GluLysMetGluAlaGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 860
Db 4695 GGAAGGCCAAGTTTGTGTCAAAATTAAGAAAGAAAGATACGGTCTTGAAGAGTTAGCTCT 4754
Qy 860 lSerGluGluLeuSerLysThrLeuValHisThrValSerValAlaValLysAspGluTh 880
Db 4755 TCTTCAAGAGAGAGAGCAAACTATTAACAGAAATGACAGGTTTATTGGAATCA 4814
Qy 880 rArgAlaValThrSerValGluGluArgSerProSerTyrLysSerAlaSerValThrGlu 900
Db 4815 GAGTCTCAGCAGCTCTCTGAA-----AGTCTAAACCTAGCTAAGGGGCTTACAG 4868
Qy 900 uProLeuGluHisThrAlaGluGluAlaMetProProValAlaGluGluValThrGluLysAs 920
Db 4869 AGACAGAGAAAGATTAAGTAAGAAATTT-----GAATCTTGAAGAAATCTTC 4913
Qy 920 pLysLeuAlaGluGluThrProValLeuThrGluThrLeuProGluGluLysAspAlaHis 940
Db 4914 TAAAGTTGCAAGAAAGTACT-----GAGTGCAGAGAGAAACA 4949

```


NUMBER OF SEQ ID NOS: 3950
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 3423
LENGTH: 4840
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Genbank Accession No. US20020142981A1 U76366
US-09-880-107-3423

Alignment Scores:

Pred. No.:	4,21e-19	Length:	4840
Score:	398.50	Matches:	382
Percent Similarity:	34.23%	Conservative:	242
Best Local Similarity:	20.95%	Mismatches:	700
Query Match:	4.94%	Indels:	505
DB:	10	Gaps:	78

US-09-902-432-4 (1-1596) x US-09-880-107-3423 (1-4840)

Qy 10 ArgSerProGluInProAlaGlySerAspThrProSerGluLeuValLeuSerGlyHis 29
Db 42 CGTCGAGGTAGCGGCGCGCGCGGCT-----GCGCGGTAT 77
Qy 30 GlyProAlaAlaGluAlaSerGlyAlaAlaGlyAspProAlaAspAlaAspProAlaThr 49
Db 78 GCGCCAGGCGCAGAGAGCGGCGGAGCTACTTCC-----CCTGATCTTACCACCATCTGCT 131
Qy 50 -LysLeuProGluInLysAspGlyGluLeuSerSerValAsnGlyValAlaGluInGlyAs 69
Db 132 GCGGCGCT-----GCGTATGTGCGTGGCGCGCGGAGGTAGAGAGCAGAGCGG 179
Qy 69 P-----ValHisValGlnG 74
Db 180 CCAGAGGTCTTCTCGGCTCAGCCCGTAACTCTTGACATCTATACACCTGGCGAACA 239
Qy 74 GlnLysAsnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 94
Db 240 AACCTCAGAGCTTGGTTCGAGACCGGAG--GCAGAGAGATGCGGCACTGCA----- 291
Qy 94 GlnLysAspValArgLysAspArgValGlnGluMetAlaAlaAsnSerThrAlaValG 114
Db 232 -----GCTAAGAAACCCTGTGTGCAGACCCCATCAGACACTCGAAGCTCGGA 341
Qy 114 AspIleThrLysAspGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 134
Db 342 AGAG--GAGGAGAAAGCAGAAAGCGAAACCGCAAGCCACCCCAAGCTGCACTCTAC 398
Qy 134 GlnLysAsnVal-----GlnGluMetValGlnProAlaGluSerG 148
Db 399 CAACCTCAGATCTGGGGGCGGAGCTTGCATCAAGCATGAAGAAAGAAAGCCAAAGCACA 458
Qy 148 n-----AlaAsnAspValGlnGlyPheLysValPheLysPheVal 161
Db 459 GACAGAGAAAGCTGGCAGACCTGGGAAATTCATGCCACACCTCGCACTGGGAAAGCGGT 518
Qy 161 GlnLysPheLysPheThrValLysLysAspLysAsnGlnLysSerAspThrValGlnLeu 181
Db 519 GCGCAACCTTCTTCTGGGAAAGTCCCGAGAGTCCAGAGAGCCCTCAGCAAAATCTAC 578
Qy 181 uThrValLysLysAspGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 201
Db 579 GTTGTCTCAGAAACTGAGAGAGAGGCGAGCGTCCCGCTTTGGAGCTGCT-----GC 632
Qy 201 nGlnProSerValGlnThrAlaValGlnLysSerAlaSerLysGlnSerGluLeuLysG 221
Db 633 CAAGCTGGAGATGCTGTCAGCG--GGCCAGGCCACACCTCCAGCAGAGACCTCCAG 689
Qy 221 nSerThrGlnLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 241
Db 690 CTCACAGTATGAGACAGAGCTGAGAGTAAAGGCTCTGAAAAAATCTCTCAGAGTCT----- 744
Qy 241 uGlnAlaGlnSerAspGlnAlaAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnG 261

Db 745 -----AGAGCTGCTCAGCCCTCGCCAAAGGG-----ACCC 776
Qy 261 nGlnLysGlnProThrLysSerProGluSerProSerSerProValAsnSerGlnThr 281
Db 777 TGGGAAAGGGCTACCCACAGACCCCTCGGAGAGCAGAGGCTGTAGCTCCAGACCA 836
Qy 281 rSerSerPheLysLysPhePheThrHisGlyTrpAlaGlyTrpArgLysLysThrSerP 301
Db 837 GCGAGG----- 843
Qy 301 eLysLysSerLysGlnAspAspLeuGlnThrAlaGlnLysArgLysGlnGlnAlaG 321
Db 844 -----AAGCCAGAGAGACTCAGA 863
Qy 321 uLysValAspGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 340
Db 864 GAGCAGCAGCAGAGAGGTATCTGAC-----AGTAGAGAGAGACCCGACG 908
Qy 341 -AlaGlnAspThrAspGlnAlaArgLeuSerAlaAspTyrGlnLysValGluLeuPro 360
Db 909 TGCCAAAGCCCTGCTTCAGCGCAGAGGCTCAGAAA----- 945
Qy 360 GlnLysAspGlnValGlnAspLeuGlnAlaSerSerGlnGlu----- 373
Db 946 -ACCTTCAAGTTCGAGCTGCTCAGACCTTGCAGAGAGTCCCCAGAAAGAGCTGC 1004
Qy 374 -----LysCysAlaProLeuAlaThrGlnValPheAspGlnLysMetG 388
Db 1005 CCGAGCGCCCTCGGAGAGAGAGGCTGAGCTGCCAAGGCCAGCGGGGAGAGCGGGA 1064
Qy 388 uAlaHisGlnGlnValValAlaGlnValHisValSerThrValGlnGlnGlnGln 408
Db 1065 GAGGAGCTCGCAGAGCAGCAGCAGAG-----GAATCGACAGTGAAGA 1106
Qy 408 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 428
Db 1107 GAGGCGCTGCTCAGCGGAAAGCTTCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1166
Qy 428 uPro-----ProGlnLysLeuAlaGlnProGlnGlnValProGlnGlnAlaG 444
Db 1167 CCTTCCCAAGAGTCCCGCAGAGAGAGGCTGCCACACT---CCTAGAGAAACAGG 1223
Qy 444 uProAlaGlnGluMetLysSerArgGluMetCysValSerGlyLysAspHisThrG 464
Db 1224 GCTTCAGACCCCGCAG-----GTCCAGGTGGGAGAGCAGAGAGA 1262
Qy 464 nLeuThrAspLeuSerProGlnGlnLysThrLeuProLysHisProGlnGlnGlnGln 484
Db 1263 GACTCAAGAGAGCAGCAGCGAGAGTCAAG-----AG 1295
Qy 484 GlnValGlnMetLeuSerSerGlnGlnArgGlnLysValGln-----GlySerPr 501
Db 1296 TGACAGAGAGAGATCGGAGAGCATGATGACAGCTCAGGTGAAGCCCTTGGGAAAAAGCC 1355
Qy 501 GLeuLysLysLeuPheSerSerSerSerGlyLeuLysLysLeuSerGlyLysGlnLysG 521
Db 1356 CCAAGTGAACCTGCTTACATGAGGAGTGGAGGCTTTG-----GG 1397
Qy 521 yLysArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 541
Db 1398 GAAA-----GCGCGCGGCGCAGTCCAGTGGGAAGGTG-----GG 1433
Qy 541 rProGlnSerAlaAspGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 558
Db 1434 GCTTCAACCCCTCAGCCAGAGTGGGAGAGTGGAGAGAGACTCAGAGACAGTATGTA 1493
Qy 558 uGlnProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 578
Db 1494 GGAAG--CATCAGACAGAGATGAGAGAGTCCCAAGCTGTGTCGCCCGCTCAGAGA 1551
Qy 578 uAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 598

Db 1552 AGTCTCTGGGAACATCTCTCC-----AGGCCAAACCACCTCCA----- 1591
Qy 598 aSerPheLysMetValThrProLysLysArgValArg-----ProSe 614
Db 1592 -----GTCCTGCAAGGGCCCCCTCAGAAAGCAGGCGCTGTAGCGTC 1635
Qy 614 rGluSerAspLysGluGluLeuLysValLysSerAlaThrLeuSerThrAs 634
Db 1636 CAGGTCAG--GCTGAAGCCCATGGCAACTCGAGAGCAGGAG--GAGTCGTGGA 1691
Qy 634 pSerThrValSerGluMetGlnAspGluValLysThrValGlyGlu----- 649
Db 1692 CAGTCGGCAGTCAGGAGCACCAGCAGCATGCTGCTGACCTCAGGCANAAACAGCTCT 1751
Qy 650 -----GluGlnLysProGluGluProLysArgArgValAspThrSerValSerTrpG1 667
Db 1752 GAAATTCCTCAGACCAAGCGCTGCCCAAGAAACCAATACCACTGCATCTGCCAAGGT 1811
Qy 667 uAlaLeuIleCysValGlySerLysLysArgAlaArgLysAlaSerSerAspAs 687
Db 1812 CGCCCTGTGGAGTGGGACCCAAACCCCGGAAGAGGAACTCGGACTTCTCCAGC 1871
Qy 687 pGluGlyGlyProArgThrLeuGlyGlyAspSerHisArgAlaGluGluAlaSerLysAs 707
Db 1872 AGGCTCATCCCGCTGTGCTGGGGCACCAGAGACCCAGAGCAGAGGATTTCTCAAGCAG 1931
Qy 707 pLysGluAlaGlyThrAspAlaValPro-----AlaSerThrGlnGluGlnAspG1 724
Db 1932 TGAGGAATCAGATAGTAGGAAGAGACAGAGGTCTTGCAGTAACCCGTGGCAGAGGCAA 1991
Qy 724 nAlaGlnGlySerSerProGluProAlaGlySer-----ProSerGluG1 740
Db 1992 GTCTGTGGGAAGCGCTCCAGGTGAAGCAGCCTCAGTCCCTGTCGAAGGGCTCTTGGG 2051
Qy 740 yGluGlyValSerThrTrpGluSerPheLysArgLeuValThrProArgLys----- 757
Db 2052 GCAAGGACTGCTCCA-----GTACTCCTGGGAAGACGGGGCC 2090
Qy 758 -----LysSerLysSerLysLeuGluGluLysAlaGluAspSerSerValGluG1 774
Db 2091 TACAGTCACCCAGGTGAAGCTGAAAGCAGGAAGACTCTGAGAGCAGTGAAGGGAATC 2150
Qy 774 nLeuSerThrGluIleGluProSerArgGluGluSerTrpValSerIleLysLysPheI1 794
Db 2151 AGACAGTGAAGAGCAGCTCATCTCCAGCACAGGTGAAACCTCAGTAAGAAACCCCA 2210
Qy 794 eProGlyArgArgLysLysArgAlaAspGlyLysGlnGluAlaThrValGluAspSe 814
Db 2211 GGCCAAAGCCCAACCCAGCTCCGCCAGAGCACCTTCAGCAAAAGGGACAAT----- 2262
Qy 814 rGlyProValGluIleAsnGluAspAspProAsnValProAlaValValProLeuSerG1 834
Db 2263 -----TCAGCCCTCGGAAAAGTTGTCATCGCAGC 2291
Qy 834 uTyrAsnAlaValGluArgGluLysMetGluAlaGlnGlyAsnThrGluLeuProGlnLe 854
Db 2292 TGCTCAAGCCCAAGCAGAGGTCTCCATCCAAAGTGAAGCCACCAAGTGAAGAACCCCA-- 2349
Qy 854 uLeuGlyAlaValTyValSerGluGluLeuSerLys-----ThrLeuValHisThrVa 872
Db 2350 -----AACAGTACCCTGTTGGGAGGGGGCCAGCATCTGTGCCATCTGT 2393
Qy 872 lSerValAlaValIleAspGlyThrArgAlaValThrSerValGluGluArgSerProSe 892
Db 2394 GGGGAAGGCGTGGCTACAGCAGCTCAGGCCCCAGACAGAGCCGAGAGGAGCTCAGGGAG 2453
Qy 892 rTrpIleSerAlaSerValThrGluProLeuGluHisThrAlaGlyGluAlaMetPro-- 911
Db 2454 CAGTGAGGAGGAGTCACAGACGTGAGGAGGAGGAGGAGCGCTGGCTCAGGCGAAGCCTTC 2513
Qy 912 -----ProValGluGluValThrGluLysAs 920
Db 2514 AGGGAAGACCCACAGATCAGAGCTGCTTGGCTCTTGCCAAGGAGTCCCCAGGAAGG 2573

Qy 920 pIleIleAlaGluGluThrProValLeuThrGlnThrLeuProGluGlyLysAspAla-- 939
Db 2574 G-----GCTGCCCAACACCTCTCT---GGGAAGACAGGCGCTTCGGCTGCCAGGCAGG 2624
Qy 940 ----HisAspAspMetValThrSer-----GluValAspPheThrSerGluAlaValTh 956
Db 2625 GAAGCAGGATGATCTCAGGAGCAGCAGCAGGAGGAATCAGACAGTGTGGGAGGACCGGC 2684
Qy 956 rAlaThrGluThrSerGluAlaLeuArg----- 965
Db 2685 AGCTGTGACCTCTGCCCGAGGTGATTAAACCCCCCTGATTTTGTGCAACCTAATCGTAG 2744
Qy 965 ----- 965
Db 2745 TCAGCTGGCCAGCTGCTACACCCGACAAAGCCAGGCTGCAAGCACCCCGAGGAAGGC 2804
Qy 966 -----ThrGluGluValThrGluAlaSerGlyAlaGlu---GluThrThrAspMetVa 982
Db 2805 CCAGGCTCGAGAGCAGCAGCAGGAGCTCTCTCCGAGAGCAGGATGAGGACGTGAT 2864
Qy 982 lSerAlaValSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValG1 1002
Db 2865 CCCCCTACACATGCTTGACTCTCTGTCATCAGAACCAATGTGTGTGACCATGCCCCTGC 2924
Qy 1002 nGluValGlu-----SerGlyValLeuAspThrGluGluGluG1 1015
Db 2925 CCACCAAGAATAAGCCCCCAAGCCAGCATGGTGGGGCCAGCAGCAAGGAGTCCAG 2984
Qy 1015 uArg----- 1016
Db 2985 TCAGATATCAGATGGCAAGAACAGAGGAGCACCGACCTCAGGTGTCAAAGAAAGAACCC 3044
Qy 1017 -----GlnThrGlnAlaIleLeuGlnAlaValAlaAspLysValLysGlu-- 1031
Db 3045 AGCTTCCCTCCCACTGACCCAGGCTGCCCTGAAGSTCCTCGCCAGAAAGCCAGTGAGGC 3104
Qy 1032 -----GluSerGlnVal----- 1035
Db 3105 TCAGCCTCTGTTGCCAGGACCCAGCCTTCAAGTGGGGTTGACAGTCTGTGGAAACACT 3164
Qy 1036 -----ProAlaThrGlnThrValGlnArgThrGly----- 1045
Db 3165 CCTGCAACAAGTCCCCAGAGCACCTCCGTCCAGGCCAAAGGACCAACAAAGTCAAGAA 3224
Qy 1046 -----SerLysAlaLeuGluLysValGluGluValG1 1056
Db 3225 ACCTAAGCTTCTGAGGTCCAGCAGGCCACCAAGGCCCTGAGAGCTCAGATGACAGTGA 3284
Qy 1056 uGluAspSerGluValLeuAlaSerGluLysGluLysAspValMetProLysGlyProVa 1076
Db 3285 GGACAGCAGCAGACAGTTCTCAGGGAGTGAAGAAATGTTGAAGGGCCCGCAGGG-- 3339
Qy 1076 lGlnGluAlaGlyAlaGluHisLeuAlaGlnGlySerGluThrGlyGlnAlaThrProG1 1096
Db 3340 -----GCCAAGTCAGCCCAACAGCTGGTCCCCCCCCCTC 3374
Qy 1096 uSerLeuGlu-----ValProGluValThrAlaAspVal-----AspHisValAl 1111
Db 3375 CAGGACAGACCCCTGTTGGAGAGACCCGACGAGAGTCCAGCAGGATGATGTGTGGC 3434
Qy 1111 aThrCysGlnValIleLysLeuGlnGlnLeuMetGluGlnAlaValAlaProGluSerSe 1131
Db 3435 GCCATCCAGTCTCTC---CTCTCAGTTATATGACCCCTGGACTAAACCCAGCCAATTC 3491
Qy 1131 rGluThr-----LeuThrAspSerGluThrAsnGlySerThrProLeuAl 1146
Db 3492 CCAGGCTCAAAGCCACTCCCAAGTAGATTCCAGCCCTCAGTTCTCTCTACTCTGGC 3551
Qy 1146 aAspSerAspThrAlaAspGlyThrGlnGlnAspGluThrIleAspSerGlnAspSerLy 1166
Db 3552 CGCCAAAGATGACCCAGATGGCAAGCAGGAGGCAAG-----CCCCA 3593


```

OY 1166 sAlaThrAlaAlaValArgGlnSerGlnValThrGluGluGlnAlaAlaThrAlaGlnLy 1186
Db 3594 ACAGCAGCAGCAGCAGTGTCTCCCTAATAACAGGTGAAAGAGGCTGCTCA----- 3645
OY 1186 sGluGluProSerThrLeuProAsnAsnValProAlaGlnGluGlnHisGlyGluGluPr 1206
Db 3646 -----GGCACCACACCTTCA-GAACTCCCGAAGCCCAAGAAAGGCGCTGGGAAC 3694
OY 1206 oGlyArgAspValLeuGluPro-----ThrGlnGlnLeuThrAlaAlaAlaValPr 1224
Db 3695 C-----CCAAGCTCAACCTGGCGCTGCAAGCAACATCACCAAGTCTCTGGGCGC 3748
OY 1224 oValLeuAlaLys-ThrGluValGlyGlnGluGlyGluValAspTrpLeuAspGlyGlu 1244
Db 3749 AACCTGGCGCTGAAATGAGGCCCAAGTGCAGGCTCACTGTGTAAGGCTCTGACTGACC 3808
OY 1244 yValValGluGlnGlnGlnValAlaPheValHisSerGlyProAsnSerGlnValAlaAla 1264
Db 3809 TCCTGGAACAGAAAGAAAGAAAGAGTGTGGACACCAAGAGAGACAGACAGAAAGGCT 3868
OY 1264 sPvalThrTrpAspSerGluValMetGly--ValAlaGlyCysGlnGluLysGluSerT 1283
Db 3869 GGGAGAGCGG--CAAGCGGAAGCTATCGGAGACCAAGCCAGTGCAGACCCCGAGAGCC 3927
OY 1283 hrGluValGln-SerLeuSerLeu--GluGluGlyGluMetGluThrAspValGluLys 1301
Db 3928 AAGAAGAAAGAAAGAGCTGGGAGCGCGGGAAGGTGGGAGGCTCTGTTTCCCAAGAAAG 3987
OY 1302 GluLysArgGluThrLysProGluGlnValSerGluGluGlyGluGlnGluThrAlaAla 1321
Db 3988 ACCGCCAGCAGCTTCCAAAGGGAAGAAAGCAAGAGAGAGCAAGAGTGTATGTCAAGAG 4047
OY 1322 ProGluHisGluGlyThrTyGlyLysProValLeuThrLeuAspMetProSerSerGlu 1341
Db 4048 AAGAAAGGAAGAGGCTCTTGGCTCCCAAGG--GCCAAGGACAGAGCCAGAAAGAGAG 4104
OY 1342 ArgGlyLysAlaLeuGlySerLeu--GlyGlySerProSerLeuProAspGlnAspLys 1360
Db 4105 CTTCAAGAAAGGAGTGGGAGCGGTGTGAAGGTGAGATCAAAAGCAACCAAGAGCAAGAG 4164
OY 1361 AlaGlyCysAlaGluValAlaGlnSerLeuAspThrThrValThrGlnThrAlaGln 1380
Db 4165 -----GAGAAAGAAATCCGACAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 4206
OY 1381 AlaValGluLysValIleGluThrValAlaIleSerGluThrGlyLysSerProGluLys 1400
Db 4207 AAGAAGAAAGAAAGAAAGAAAGAAAGGCTCAACCAAGATTCGATGCACCG----- 4260
OY 1401 ValGlyAlaHisLeuLeuProAlaGluLysSerSerAla----- 1413
Db 4261 ---TCCCAAGAAAGAAAGAAAGAAAGAAAGAAAGAGAGAGAGAGCTGTATGACAGCA 4317
OY 1414 -----ThrGlyGlyHisTrpThrLeuGlnHisAla 1423
Db 4318 CCAAGCAAGAGAGAGAGAGATTCCTTACCGAGAGAGTGGCCAT-----CCCAATGCC 4368
OY 1424 GluAsp-----ThrValProLeu 1429
Db 4369 TCTGACCTCCACCGACCTGTGCCACATGGTGGAGCTAAACCTGTACCTTCCCTCGC 4428
OY 1430 GlyProGlu-----SerGlnAlaGluSerIleProIleIleValThrProAlaProGlu 1447
Db 4429 TCCACAGAGAGAGAGAGAGCTTCCAGAGG-GTCCCTGTGTGTCGCAAGCCAGAGT----- 4481
OY 1448 SerThrLeuHisProAspLeuGlnGlnGlnGlnLeuSerAlaSerGlnArgGluAspSerGlu 1467
Db 4482 -----AGCTTGGCGGAGAGGCTGTCTC-----AAGGAGAAAGTGGAC 4517
OY 1468 GluGluAspLysProAspAlaGlyProAspAlaAspGlyLysGluSerThrAlaIleGlu 1487
Db 4518 CAGCTCCCATGACCTCAACCCCACTCCCAACAGAGAGAGCTTCATATGATGTGTACAG 4577
OY 1488 LysValLeuLysAlaGlu---ProGluIleLeuGluLeuGlnLysSerAsnLysIle 1506

```

```

Db 4578 TATATGATTTTTTTTAAGTACCTCTCTCTT----- 4610
OY 1507 ValLeuAsnValIleGlnThrAlaValAspGlnPheAlaArgThrGluThrAlaProGlu 1526
Db 4611 -----CCACAG 4616
OY 1527 Thr--HisAlaTyTrpAspSerGlnThrGlnValProAlaCysArgLeuAspSerArgGlu 1545
Db 4617 ACCCCACATGCCAAAGGCTCGGAGACTTCCACACACTTCTCCACAGATCCAGCTAAG 4676
OY 1546 ProAsnArgCysTrpThrLysMetLysAspAlaLysMetLysHisProValProGluPro 1565
Db 4677 CCGACCTGTGCTCT-----CATCCCGTCCG----- 4702
OY 1566 ArgGluAspLeuGlnValLeuThrValLeuGlnAlaTrpAlaGlnProArgLysCysLeu 1585
Db 4703 -----CTCGGTCTGTGCGATCCGAGGCTTGTCTT 4735
OY 1586 PRO 1586
Db 4736 CCT 4738

RESULT 15
US-09-880-107-3311
; Sequence 3311, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3311
; LENGTH: 11185
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 U16306
US-09-880-107-3311

Alignment Scores:
Pred. No.: 1,98e-18 Length: 11185
Score: 395.00 Matches: 431
Percent Similarity: 31.56% Conservative: 252
Best Local Similarity: 19.92% Mismatches: 725
Query Match: 4.89% Indels: 756
DB: 10 Gaps: 102

US-09-902-432-4 (1-1596) x US-09-880-107-3311 (1-11185)
OY 25 ValLeuSerGlyHisGlyPro---AlaAlaGluAlaSerGlyAlaAlaGlyAspProAla 43
Db 2622 GGGCTTTGGCCCATGATGATCTTAAAGTGTGAAGCAGCAGCTGTATCAAAATGTCATGG 2681
OY 44 AspAlaAspProAlaThr-----LysAsnGlyGlnLeuSerSerValAsnGly 50
Db 2682 GATGAAGATATATCAATCAATCCAAAGCTTTAAGTCTACAGAACTTCAAGCTTTCAGAAA 2741
OY 51 LeuProGln-----LysAsnGlyGlnLeuSerSerValAsnGly 63
Db 2742 TTGCCCCCTGCTTACTCAACAATGTGGGATGATGTGAAGAAAGATAAAGACATCCCAAGT 2801
OY 64 ValAlaGluGlnGly-----AspValHisValGlnGluGlnGlnGlnGlnGlnGln 81

```

Db 2802 TTCACTGAAGATGGACAGATGAATTTACTCTTATTTCCAGATAGTACTCAAAAGCAGTTA 2861
 QY 82 GluGluValValAspGluAspValGlyGlnArgGluSerGluAspValArgGluLysAsp 101
 Db 2862 GAGGAGGTACTGATGAAGACATAGACGCCCATGGAAAATTCACAATTAGATTTACGCCA 2921
 QY 102 ArgValGluGluMetAlaAlaAsnSerThrAlaValGluAspIleThrLysAspGlyGln 121
 Db 2922 ACTACATCAACTGGTATTCGAGAAAAGTCAACTTTGAGAGATTCTACA----- 2969
 QY 122 GluGluThrSerGluIleIleGluGlnIleProAlaSerGluAsnAsnVal---GluGlu 140
 Db 2970 -----ACTGAAGAAAAGTTCCACCTATCAAGCACTCAAGCCAGTTTATGCAACC 3023
 QY 141 MetValGlnProAlaGluSerGlnAlaAsnAspValGlyPheLysLysValPheLysPhe 160
 Db 3024 ATGAAGGAAGTCTTGGGTGAAGTAGAAGATGTGGACCTCTCTAAGCCAGTATCTACT 3083
 QY 161 Val-----GlyPheLysPheThrValLysLysAsp 170
 Db 3084 GTTCCCCAAATTCACACACTTCAGAGGTGGAAGGATTAGCATTTGTTAGTTATAGTAGC 3143
 QY 171 LysAsnGlu-----LysSerAspThrValGlnLeuLeuThrValLys 184
 Db 3144 ACCAAGAGCCTACTACTTATGTAGACTCTTCCCATACCACTCTCTTCTGTAATTTCCC 3203
 QY 185 LysAspGluGlyGluGlyAlaGluAlaSerValGlyAlaGlyAspHis-----GlnGlu 202
 Db 3204 AAGACAGACTGGGAGGTGTAGTACCTTCTGTTCCATCAGAAGATGAAAGTTCTAGGTGAA 3263
 QY 203 ProSerValGluThrAlaValGlyGluSerAlaSerLysGluSer-----Glu 218
 Db 3264 CCTCTCAAGACATACTTGTTCATTGATTCAGACTCGCCTTGAAGCGACTATTCTCCAGAA 3323
 QY 219 LeuLysGlnSerThrGluLysGlnGlyThrLeuLysGlnGlu-----Gln 234
 Db 3324 ACTATGAGAAACAATAATCAGAGGGGAACAATCAGGAAGAATTCCTTGGAAAGAA 3383
 QY 235 SerSerThrGluIleProLeuGlnAla----- 243
 Db 3384 CAGACTGCAGAGAAACAGTCTCGTCTCTCAGTCTTACAGTTGGACTCCCAAGGAGCA 3443
 QY 244 -----GluSerAspGlnAlaAlaGluGluAlaLysAspGlu 256
 Db 3444 GTAAACCACTGGATGAACAAGAGGGCGATGGATCAGCATATACAGTCTCTGAAGATGAA 3503
 QY 257 -----GlyGluGlu----- 259
 Db 3504 TTGTTGACAGGTTCTGAGAGGGTCCAGTTTATAGAAACAACCTCCAGTTGGAAAATTTGAT 3563
 QY 260 -----LysGlnGluLysGlu 264
 Db 3564 CACAGTGTGTTATCCACAGGTGTGTAACTGAGCACAAGAGTGAACAGATGAAGTG 3623
 QY 265 ProThrLysSerProGlu-----SerProSerSerProValAsn 277
 Db 3624 GTAAACATACACACCGCATTTGGGCCAAAAGTATCTTTAAGTCCAGGCGCTGAACAAAA 3683
 QY 278 SerGluThr-----ThrSerSerPheLysLysPhePheThr 289
 Db 3684 TATGAACAGAAAGGTAGTAGTACAACAGGATTTTACATCATCTTTGAGTCTCTTTAGTACC 3743
 QY 290 HisGlyTriAlaGlyTriArgLysLysThrSerPheLysLysSerLysGluAsp---Asp 308
 Db 3744 CACATTACCAGCTTATGGAGAAACCACTACTGAGAAACATCCCTAGAGGATTTGAT 3803
 QY 309 LeuGluThrAlaGluLysArgLysGluGlnGluAlaGluLysValAspGluGluLys 328
 Db 3804 TTAGGTCAGGATTATTTGAAAAGCCCAAGCCACAGCACTCATAGAAATTTTCAACAATC 3863
 QY 329 GluLysThrGluProAlaSer----- 335

Db 3864 AAAGTCACAGTTCCAAGTGTATATTACCACCTGCTTCAGTTTCAGTACAGACTTCACACA 3923
 QY 336 -----GluGluGlnGlu 339
 Db 3924 ACTTCAGCATTTCAAGCCATCTTCGCGATCACTTAAGAAAACCACTCTCATCGACAGGAA 3983
 QY 340 ProAlaGluAspThrAspGlnAlaArgLeuSerAlaAspTyArgLysValGluLeuPro 359
 Db 3984 CCTGGTGAAGAAACAACCAAGTGCATGTTAATCATTTGGAGAATCAACATCTCTCATGTTCT 4043
 QY 360 -----LeuGluAspGlnValGly-----AspLeuGluAlaSerSer 371
 Db 4044 CCCACTACCTTTGAAGATATTGTAGCCAAGGAACAACCGATATTGTAGAGAGTAT 4103
 QY 372 GluGluLysCysAlaProLeuAlaThrGlu-----ValPheAspGlu 385
 Db 4104 TTCACGACTTCAAGTCTCTCTGTACACAGCAACAAGACCCCACTGTGGAAGACAAA 4163
 QY 386 LysMetGluAlaHisGlnGluValValAla----- 395
 Db 4164 GAGGCCCTTTGGACCTCAGCGCTTCTACGCCACAGCCCCCAGCAAGCAACAAATTTTAC 4223
 QY 396 ---GluValHisValSerThrValGluLysThrGluGluGluGlnGlyGlyGlyGlu 414
 Db 4224 CCTGACATTAAATGTTTATATTATGAGTGCAGAGAAAATAAGACAGTCAATGAGTGAT 4283
 QY 415 AlaGluGlyGlyValValValGluGlyThrGlyGluSerLeuProGluLysLeuAla 434
 Db 4284 -----TTGAGTGTAAATTTGGTCACTCAATAGATTCAAGATCT---AAA 4322
 QY 435 GluProGlnGluValProGlnGluAlaGluProAlaGluGluLeuMet----- 450
 Db 4323 GAAGATGAACCTTTGTAGTGAAGAAACAGATCCAGTGCATGATCTAATGGCTGAATTTTA 4382
 QY 451 -----LysSer 452
 Db 4383 CCTGAATTCCTGCACATAATTGAATAGACCTATACCACAGTGAAGAAAATGAAGAGAA 4442
 QY 453 ArgGluMetCysValSerGlyGlyAspHisThr-----GlnLeuThrAsp 467
 Db 4443 GAAGAAGAGTGTGCAAAATGCTACTGATGTGACAAACCAACCCCATCTGTGCAGTACATAAAT 4502
 QY 468 LeuSerProGluGluLysThrLeuProLysHisProGluGly----- 481
 Db 4503 GGAAGACATCTCGTTACCACTGTGCCCAAGGACCCAGAGAGCTGCAGAAGCTAGCGCTGGC 4562
 QY 481 ----- 481
 Db 4563 CAGTTTGAAGTGTTCGACCTTCTCAGAAATTTCTCGGACAGCTCTGAAAGTGATCTCAT 4622
 QY 482 -----IleValSerGluValGluMet-----LeuSer 490
 Db 4623 CCATTTGTAATAGCCAAAACGGNAATTTGCTACTGTGTGCAACCTAATGAATCTACAGAA 4682
 QY 491 SerGlnGluArgIleLysValGlnGlySer----- 500
 Db 4683 ACAACTGAGTCTCTTGAAGTTATCGAAGCTGAGACTTACCTGAAACATCAGAAACAT 4742
 QY 501 -----ProLeuLysLysLeuPheSerSer 508
 Db 4743 TTTTCAGTGGTGAGCCTGATGTTTCCCAAGTCCCATTCATGAGGAATTTGAAGT 4802
 QY 509 SerGlyLeuLysLysLeuSerGlyLysLysGlnLysGlyArgGlyGlyGlyGlyAsp 528
 Db 4803 GGNACAGCCAAAAA-----GGGCAGAAATCAGTCACAGAGAGA-----GAT 4844
 QY 529 GluGluProGlyGluTyArgLysHisIleHisThrGlu----- 540
 Db 4845 ACTGAAGTTGGTTCATCAGGCACATGAACATACATACTGAACCTGTATCTCTGTTTCTCGAAGAG 4904
 QY 541 ---SerProGluSerAlaAspGluGlnLysGlyLysSerAla----- 554
 Db 4905 TCTTCAGGAGAGATTGCCATTGACCAAGAAATCTCAGAAAATAGCCTTTGCAAGGGGTACA 4964

QY 555 -----SerSerProGluGluProGluGluThrThrCysLeuGluLysGlyProLeuGlu 572
 Db 4965 GAAGTAACATTGTTGGTGAAGTGAAGAAAAAGTACTTCTGTACATACATACCTCCACTATA 5024
 QY 573 AlaProGlnAspGlyGluAla-----GluGluGlyThrThrSerAspGlyGlu 588
 Db 5025 GTTCCAAAGTTTCGATCAGCATATGTTTTCAGAGAGAGAAAGCACTTCCCTAATAGA--- 5081
 QY 589 LysLysArgGluGlyIleThrProTrp-----AlaSerPheLysMet--- 603
 Db 5082 -----AATCTTGCCAGATGACCTGTTGCTTCCAAAGAAAGCTGG 5123
 QY 604 -----ValThrProLysLysArgVal----- 610
 Db 5124 GTAGAAGCAACTCTTAGACAAGTTGTAGAGCTCTCAGGAGTCTTCGATTCGCAATTACA 5183
 QY 611 ArgArgProSerGluSerAspLysGluGluLeuGluLysValLysSerAlaThrLeu 630
 Db 5184 GAAGGCTCTGGAGAGACAGAAAGATGAAGATACATGTTACCCATGGTAAGTATT 5243
 QY 631 SerSerThrAspSerThr-----ValSerGlu 639
 Db 5244 TCACAGAGAAATACCTCTGATACACTTACTTAGACACTAGACAGATTAATACAGAA 5303
 QY 640 MetGlnAspGluValLysThrValGlyGluGluLysProGluProLysArgArg 659
 Db 5304 ACCTTTTGGAGTTCCTGCAACACCATTTATCCAGTTTCTGCAACACCTTCTGCAGAA 5363
 QY 660 Val-----AspThrSerValSerTrpGluAlaLeuIle 670
 Db 5364 GTGGGCGCTTACCAAGTTTGTAACTGAAGAAACAGACACTTCT---GAGTGGATTTCAGTACC 5420
 QY 671 CysValGlySerSerLysLysArgAlaArgLysAlaSerSerSerAspGluGlyGly 690
 Db 5421 ACTGTTGAGAGAAAGAAAGAAAGAG---GAGGAGGAGGAACT 5459
 QY 691 ProArgThrLeuGlyGlyAspSerHisArgAlaGluGluAlaSerLysAspLysGluAla 710
 Db 5460 ACAGGTACG-----GCT 5471
 QY 711 GlyThrAspAlaValProAlaSerThrGluGluGlnAspGln-----AlaGlnGly 727
 Db 5472 TCTCACTTTGAGGTATTCATCTACACAGACGTCGATCAATTAATTTTACCTTGA 5531
 QY 728 SerSerSerProGluProAlaGlySerProSerGluGlyGluGlyValSerThrTrpGlu 747
 Db 5532 TTAGAAAGTCCAAATGTAGTACATCTAGATTCAGGT-----ACCAGGAAA 5579
 QY 748 SerPheLysArgLeuValThrProArgLysLysSerLysSerLysLeuGluGlyLys--- 766
 Db 5580 AGTTTATATGTCCTTGAACAACACACAGTCTGAAAGGAAATGACAGATTCTACTCCT 5639
 QY 767 -----AlaGluAspSerSerValGluGluLeuSerThrGluIleGluProSerArgGlu 784
 Db 5640 GTCTTACAGAAACAATACATTAGAAATTTGGGGCAGG-----ACCACCTGAG 5690
 QY 785 GluSerTrpValSerIleLysLysPheIleProGlyArgArgLysArgAlaAspGly 804
 Db 5691 CACAGACGATC-----CATCAACCTGGG 5714
 QY 805 LysGluGluGlnAlaThrValGluAspSerGlyProValGluLeuAspGluAsp----- 822
 Db 5715 GTTCAGAGAGGCGCTGACCTCTCCACGTAAGTCCTCTCTTATGAGAGCAGGCG 5774
 QY 823 -----AspProAsnValProAlaValValProLeuSer 833
 Db 5775 TCTGAGAAAGCTGCTGCCGACCCAGAAACCAACCACTGTTTCTTCATTTTAAACGTA 5834
 QY 834 GluTrpAsnAlaValGluArgGlyLys-----MetGlu 844
 Db 5835 GAGTAT---GCAATTCAAGCCGAAAGAAAGTAGTGCCTGCACTTTGTCTCCGCACTGGGAA 5891

QY 845 AlaGlnGlyAsnThrGlu----- 850
 Db 5892 ACTACATTCTCCACTGACGCCAACAGAGCTGTTTGGATACAGTAATGACAGAGTACT 5951
 QY 851 -----LeuProGlnLeuLeuGlyAlaValTrpValSerGluGluLeuSerLys--- 866
 Db 5952 GCTGAAATATTAACCCAAACATCCACGGAATATGATGATTTTACAGCCGCTTAGAGAAACCA 6011
 QY 866 ----- 866
 Db 6012 AATTATGGGAGAAATAAGGGGCTTTTCCACAGGTTTCTTTGAGAGAAATTCAGT 6071
 QY 867 -----ThrLeuValHisThrValSer-----ValAlaVal 876
 Db 6072 GGTCACCTTAGAATTAACATCAACGTGCTTATCCCATAGCAAAAGAAACGGTAATG 6131
 QY 877 IleAspGlyThr---ArgAlaValThrSerValGluGluArgSerProSerTrpIleSer 895
 Db 6132 ATGGAAGGCTCTGGAGATGAGATTTAGGAGACCCAGACTTACCATCTTACAGTACT 6191
 QY 896 AlaSerValThrGluProLeuGluHisThrAlaGlyGluAlaMetPro----- 911
 Db 6192 ACTTCAGTT---CACATAGTCACATATCTGACTCAGAAAGACCCAGTAGACCATG 6245
 QY 912 -----ProValGluGluValThrGlu-----Lys 919
 Db 6246 GTCAGACCTTACGCTTCCCTGGGAGAGATTTCATCTCAGCTGAGGCTCAGGTGAG 6305
 QY 920 AspIleIleAlaGluGluThrProValLeuThrGlnThrLeuProGluGlyLysAspAla 939
 Db 6306 CAACGTGCACAGTCAGACAGCTGTGTGT---CCAGTCTTCCAGTGCCTGCAAG 6362
 QY 940 HisAspAspMetValThrSerGluValAsp---PheThrSerGluAlaValThrAlaThr 958
 Db 6363 TTTTCTGTACAGCTTCTCCATTCATTCAGCAAGATTTGGAGAAAGTGGTACTGTCAAT 6422
 QY 959 Glu-----ThrSerGluAlaLeuArgThrGluGluVal-----ThrGluAlaSer 973
 Db 6423 GAAATGTATAGAAAGTCCACCTTTTACCAACAGCAAGAGTGAAGTCAAAAGCTCCA 6482
 QY 974 GlyAlaGluGluThrThrAspMetValSerAlaValSerGluLeuThrAspSerProAsp 993
 Db 6483 GTAGAGAGAGAGAGTAAGTACGTGACAGTTCA-----ACAACTTCTCCCA 6536
 QY 994 ThrThrGlu-----GluAlaThrProVal---GlnGluVal 1004
 Db 6537 ACTATAGAGCCACCAATTAATGTCTAGGCAAGAAAGTCAACCTGTAAACAAAGAAAT 6596
 QY 1005 GluSerGlyValLeuAspThrGluGluGluGluArgGlnThrGlnAlaIleLeuGlnAla 1024
 Db 6597 GAAAGTGAACA-----ACATCAGAGAGAACAAATTCAAGAAAGAGTCAATTGAATCC 6650
 QY 1025 ValAlaAspLysValLysGluGluSerGlnValProAlaThrGlnThrValGlnArgThr 1044
 Db 6651 CCTCAAACTCTCCGCAACAGAACAAACAAATCTTGATTCACACACTTAAGTAAACT 6710
 QY 1045 GlySerLys----- 1047
 Db 6711 GAACCAAAACACAGATTATTCGTACTTAACAAAGAAAACTTAGAGTGATGATAAA 6770
 QY 1047 ----- 1047
 Db 6771 GAAATGAGAGAGAAAGACACTTCTTACTTAACATGCTACTCCAGATCCAGATCAAT 6830
 QY 1048 -----AlaLeuGluLys----- 1051
 Db 6831 GCGTTGAATCTTACACAACCTCCCTGAAGCTACTGAAABATCAATTTTCTTAGCT 6890
 QY 1052 -----ValGluGluValGluGluAspSerGluValLeu 1062
 Db 6891 ACTGCATTAGTAAGTAATCTATACAGCTGAACATGTAGTCAACATTCACCAATCAAA 6950
 QY 1063 AlaSerGluLysGluLysAspValMetProLysGly-----ProValGlnGluAla 1079

[illegible]

Search completed: December 13, 2002, 06:51:00
Job time : 1298 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 13, 2002, 00:22:56 ; Search time 4124 Seconds
(without alignments)
6267.700 Million cell updates/sec

Title: US-09-902-432-4
Perfect score: 8073
Sequence: 1 MGAGSSTQRSPEQAGSDT.....AWAQRKCLPRLQLKAPVSK 1596

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgn2.1/USPTO.spool/US0902432/runat.12122002.141452.25812/app_query.fasta_1.1735
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US0902432 @CGN 1 1 3512 @runat.12122002.141452.25812 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_esti:*
10: gb_estc2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fam:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1268.5	15.7	1269	11	BC015830	BC015830 Homo sapi
2	1041.5	12.9	776	13	BG974884	BG974884 602842771
3	1000	12.4	668	10	B8546856	B8546856 B8546856
4	908	11.2	661	13	B1653823	B1653823 603298753
5	856.5	10.6	674	9	AL045541	AL045541 DKFZP434K
6	854.5	10.6	1117	14	BM04332	BM04332 AGENCOURT
7	853.5	10.6	1153	14	D88456	D88456 D88456 Mouse
8	821	10.2	610	9	AI225639	AI225639 u113009.Y
9	806.5	10.0	646	10	B8618075	B8618075 601462263
10	785.5	9.7	796	9	AI796191	AI796191 wh43905.X
11	782	9.7	873	14	BQ717711	BQ717711 AGENCOURT
12	776.5	9.6	924	14	BQ221404	BQ221404 AGENCOURT
13	748	9.3	530	9	AI156819	AI156819 u144c01.Y
14	744	9.2	679	14	BQ443215	BQ443215 UI-M-EVO-
15	713	8.8	573	12	B400579	B400579 UI-R-CAO-
16	701	8.7	584	12	B448126	B448126 hr8h12.X
17	690.5	8.6	532	10	B8350908	B8350908 ht63905.X
18	686.5	8.5	532	10	B163549	B163549 601769945
19	686	8.5	641	10	B8665546	B8665546 B8665546
20	682	8.4	875	14	BQ427917	BQ427917 AGENCOURT
21	673.5	8.3	1045	12	B8394092	B8394092 602456258
22	666.5	8.3	539	12	B835245	B835245 354322 MA
23	657	8.1	644	13	B1446672	B1446672 dac81f11.
24	656	8.1	490	9	AL041664	AL041664 DKFZP434M
25	653	8.1	467	9	AL041405	AL041405 DKFZP434C
26	637.5	7.9	911	12	BG751345	BG751345 602730292
27	632	7.8	923	14	BQ913238	BQ913238 AGENCOURT
28	598.5	7.4	556	12	B8347898	B8347898 dac81f11.
29	592.5	7.3	527	9	AL041657	AL041657 DKFZP434M
30	592	7.3	461	12	B8364956	B8364956 CM2-NN114
31	591.5	7.3	651	13	BJ044180	BJ044180 BJ044180
32	590.5	7.3	861	13	B1154270	B1154270 602903056
33	590	7.3	459	9	AL040086	AL040086 DKFZP434B
34	577.5	7.2	466	12	B859328	B859328 UI-M-AQO-
35	568.5	7.0	724	14	BM721210	BM721210 UI-E-EQO-
36	566.5	7.0	561	9	AL703803	AL703803 DKFZP686J
37	560	6.9	812	12	B8538601	B8538601 602567330
38	557	6.9	465	12	B8364993	B8364993 CM2-NN114
39	542	6.7	423	14	W82339	W82339 me98d02.r1
40	539	6.7	446	10	B8120002	B8120002 UI-R-CAO-
41	537	6.7	431	9	AI322363	AI322363 me98d02.Y
42	513	6.4	1105	14	BQ935865	BQ935865 AGENCOURT
43	506	6.3	883	13	BM047462	BM047462 603628755
44	504	6.2	333	9	AI548287	AI548287 UI-R-C3-t
45	500.5	6.2	436	9	AA159505	AA159505 z072c07.B

ALIGNMENTS

RESULT 1
LOCUS BC015830
DEFINITION Homo sapiens, clone IMAGE:4691862, mRNA.
ACCESSION BC015830
VERSION BC015830.1 GI:16198369
KEYWORDS HTC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1269)
REFERENCE Strausberg,R.
AUTHORS Direct Submission
TITLE Submitted (15-OCT-2001) National Institutes of Health, Mammalian
JOURNAL

QY 1 MetGluValaGlySerSerThrGluGlnArgSerProGluGln---ProAlaGlySerAsp 19
Db 25 ATGGAGCCGCCGGAGACTCCACCGAGCAGCGCAGCCCGGAGCAGCCGCCCGAGGGAGACTCC 84
QY 20 ThrProSerGluLeuValLeuSerGlyHisGlyProAlaAlaGluAlaSer--GlyAla 38
Db 85 ACGCGCGGTGAGCCCGAGCCGAGCGGGCGGGCGCCCTCGGCGGAGCGGGCGCCAGACAC 144
QY 39 AlaGlyAspProAlaAspAla-----AspProAlaThrIysLeuProGlnIysAsnGly 56
Db 145 ACCGCGGACCCCGGCATCGCTGCTCGGACCCCGCCACCAAGCTCCACAGAAAGAAAGT 204
QY 57 GluLeuSerSerValaSnGlyValAlaGluGlnGlyAspValHisValGlnGluGluAsn 76
Db 205 CAGCTGTCACATCCATCAATGGCGTAGCTGAGCAAGATAGACTCAGCCTCCAGAGAGGTGAC 264
QY 77 GlnGluGlyGln-----GlnGlu 82
Db 265 CTTAAATGCCCGAAAGAGAGCCCTGAACGCTGAAGAGCCCTAAACACCCAGAGGAAAGA 324
QY 83 GluValValaLeuGluAspValGlyGlnArgGluSerGluAspValArgGluIysAspArg 1020
Db 325 GAAGTCATTGTCCACAGAGAGTTGGACAGAGAAACTCTGAAGATGTGAGCAAGAAAAGAGACTCC 384
QY 103 ValGluGluMetAlaAlaAsnSerThrAlaValGluAspIleThrIysAspGlyGlnGlu 1220
Db 385 GATAAAGAGATGGCTACTCAATGCTACGCGGTGTTTCAACGCATCACAGATGATGGCCACAGAG 444
QY 123 GluThrSerGluIleIleGluGlnIleProAlaSerGluAsnAsnValGluGluMetVal 1424

Db	445	GAGACACCCGAAATTAATGCAACAGATTCCTTC	CGAGAAACCAATTTGAGAGACTAAC	504
Qy	143	GlnProAlaGluSerGlnAlaAsnAspValGly	PhelyslyValPheIysPheValGly	162
Db	505	CAACCCACTGATGCCAGGCTTAATGAATTTGA	TTTAAAGAGGCTTTAAAGTTTGTTGGC	564
Qy	163	PhelysPheThrValIlysAspIysAsnGluIys	SerAspThrValGlnLeuLeuThr	182
Db	565	TTTAAATTCACCTGCGAAAAAGATTAAGACAG	AGAAAGCTGACACTGTCCAGCTACTCACT	624
Qy	183	ValIlysIysAspGluGlyGlyGlyValAlaGlu	AlaSerValGlyAlaIysAspHisGlnGlu	202
Db	625	GTGAAGAAAGATGAAGGGAGGAGCA-----	GCAGGGCTGGCGACCAACAAGAC	675
Qy	203	ProSerValGluThrAlaValAlaGlyIleuSer	AlaSerIysGluSerGluLeuIlysglnSer	222
Db	676	CCCAAGCTT-----GGGGCTGAGAAAGCAC	ATCTCCAAAGAAACGAAACCAATCT	729
Qy	223	ThrGluIysGlnGluGlyThrLeuIysGlnGlu	IysSerThrGlnIleProLeuGln	242
Db	730	ACAGGAAACCCGAAAGAGACCCCTGAAGCTAG	CAAAACCGCAGAAATTTCTCCCCCA	789
Qy	243	AlaGluSerAspGlnAlaAlaGluGlnGluAla	IlyAspGluGlyGlyGlnIlysglnGlu	262
Db	790	GCCGAATCTGGCCCAAGCAGTg---GAGAAAT	GCAAGAGAGAGAGAGAGAAACAAGAA	846
Qy	263	IlysgluProThrIlysserProGluIysProSer	serProValAsnSerGlnThrThrser	282
Db	847	AAAGAACTTACCAAGTCTGCAGAAATCTCCGA	CTGATCCCGTGACCAAGGAAACAGATCA	906
Qy	283	SerPheIysIysPhePheThrHisGlyTyrAla	GlyTyrAlaGlyIysThrSerPheIys	302
Db	907	ACCTTCAAAAAATCTTCACCTCAAGGTTGGG	CCGGCTGGCCAAAAAGACCAAGTTTCAG	966
Qy	303	IysSerIysGluAspAspLeuGluThrAlaGlu	IysAspIysGlyGlnGlnGluAlaGluIys	322
Db	967	AAGCCGAAGAGGATGAAGTGAAGCTTCAGAA	AGAAAGAAACAGAGCCAGAAAAA	1026
Qy	323	ValAspGluGluGluIlysglyIysGlyThrGlu	ProAlaSerGlnGluIlysgly-----	337
Db	1027	GTAAGCACAGAAAGAAAGCGAAAGCAAGGTT	GCCTCCGAGAAACGTACCGCTCCGAG	1086
Qy	338	-----GlnGluProAlaGluAspPheThrAsp	GlnAlaArgLeuSerAlaAspTyr	353
Db	1087	CAAGCCCAACCAACAGAGCCGGCGAAGATGCC	ACAGACCGCCCGTTTACAGCTGAATAT	1146
Qy	354	GluIysValGluLeuProLeuGluAspGlnVal	GlyAspLeuGlnAlaSerSerGlnGlu	373
Db	1147	GAGAAAGTTGAGCTGCCCTCAGAGAGCAATCT	ACGTGCTCGCAGAGGACCTTCTGAAGAG	1206
Qy	374	IysCysAlaProLeuAlaThrGluValPheAsp	GluIys	386
Db	1207	AAACCTGCTCGTTGGCAGAGAGTGTATTATGA	AAAAA	1245

RESULT 2
BG974884

LOCUS	BC974884	776 bp	mRNA	linear	EST 12-JUN-2002
DEFINITION	602842771F1 NCI_CGAP_Mam4 Mus musculus cDNA clone IMAGE:4978744 5',				

ACCESSION	BG974884
VERSION	BG974884.1
	GI:14362521

KEYWORDS	EST.
SOURCE	house mouse.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus

REFERENCE
1 (bases 1 to 776)
AUTHORS
NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)

Email: cgapbs-romail.nh.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furcht

Ph.D.

CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM10975 row: m column: 17
 High quality sequence stop: 753.

FEATURES

source

Location/Qualifiers
 1..776
 /organism="Mus musculus"
 /strain="NMRI"
 /db_xref="taxon:10090"
 /clone="IMAGE:4978744"
 /clone_lib="NCI CGAP Mam4"
 /issue_type="tumor, gross tissue"
 /dev_stage="5 months"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
 Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
 Library constructed by Life Technologies. Investigators
 providing samples: Lothar Hennighausen/Priscilla Furth,
 NIH reference for transgenic model: Li et al., Cell Growth
 and Differentiation 7, 3-11 (1996)."
 BASE COUNT 278 a 181 c 216 g 101 t
 ORIGIN

Alignment Scores:

Pred. No.: 4,94e-75 Length: 776
 Score: 1041.50 Matches: 220
 Percent Similarity: 89.27% Conservative: 13
 Best Local Similarity: 84.29% Mismatches: 22
 Query Match: 12.90% Indels: 6
 DB: 13 Gaps: 3

US-09-902-432-4 (1-1596) x BG974884 (1-776)

QY 99 GluLyAspArgValGluGluMetAlaAlaAsnSerThrAlaValGluAspIleThrLys 118
 DB 1 GAAAGAGACCGAGTAAAGAAATGGCGACCGCAGTCCACAGTGGTTGAAGATATACAAAG 60
 QY 119 AspGlyGlnGluThrSerGluIleIleGluGlnIleProIleSerGluAnsnVal 138
 DB 61 GACGACGAGGAGAAACACCGGAAATAATCGAACAGATCCCTCTTCAGAGACCAATGTG 120
 QY 139 GluGluMetValGlnProAlaGluSerGlnAlaAsnAspValGlyPheLysValPhe 158
 DB 121 GAAGAATGGCGAGGCTGCTAGTCCCAAGCTAATGACGTCGGCTTCAAGAGGTATT 180
 QY 159 LysPheValGlyPheLysPheThrValLysAspLysAsnGluLysSerAspThrVal 178
 DB 181 AAATTTGTTGTTTAAATTCACGGTGAAGAAGATAAAACGAAAGTCAGATACCGTC 240
 QY 179 GlnLeuLeuThrValLysLysAspGluGlyGluGlyAlaGluAlaSerValGlyAlaGly 198
 DB 241 CACCTACTCACTGCAAGAGGATGAAGGCGGAAGGGGCGAGACCTCCGTCGGAGCAGGA 300
 QY 199 AspHisGlnGluProSerValGluThrAlaValGlyGluSerAlaSerLysGluSerGlu 218
 DB 301 GACCACAGAGCCCGAGTGGAGACC--GTCGGCGAATCAGATCCAAAGAAGGTAG 357
 QY 219 LeuLysGlnSerThrGluLysGlnGluGlyThrLeuLysGlnGluGlnSerSerThrGlu 238
 DB 358 CTGAAGCAATCCACAGAGAGCAGAAGGACCCCTGAGCAAGCACAGACACAGAA 417
 QY 239 IleProLeuGlnAlaGluSerAspGlnAlaAlaGluGluGluAlaLysAspGluGlyGlu 258
 DB 418 ATTCCTCCCTCAAGCCGAATCTGCTCAAGGAGCCGAGGAAGACGAGCAAAAGATGAGAA 477
 QY 259 GluLysGlnGluLysGluProThrLysSerProGluSerProSerProValAnsnSer 278
 DB 478 GAAACCGAGAGAAAGAACCTCAAGGCCCTTAGAATCTCCGAGCCCGCCCTGTCAGCAAT 537

QY 279 GluThrThrSerSerPheLysLysPhePheThrHisGlyTrpAlaGlyTrpArgLysLys 298
 DB 538 GAGACAACTCTCTCTCAAGAAATCTTCACTCACGGCTGGCGCGCTGGCGCAAGAG 597
 QY 299 ThrSerPheLysLysSerLysGluAspAspLeuGluThrAlaGluLysArgLysGluGln 318
 DB 598 ACCAGCTTCAAGAAACCAAGGAAGATGATCTGGAACCTTCCGAGAGAGACAGGAGCAA 657
 QY 319 GluAlaGluLysValAspGluGluGluLys--GluLysThrGluProAla---SerGlu 336
 DB 658 GAGGCTGAAAAAGTAGACCCGAGGAGCAAGGGGACCAAGACAGAGCCGCCGCGAG 717
 QY 337 GluGlnGlu-ProAlaGlu-AspThrAspGlnAlaArgLeuSer-AlaAspThrGlu 354
 DB 718 GAGCAGGAAGCCTGCAGAAAGGCGACAGACGAGCCAGGATGACCCAGCCGACTATGAG 774
 RESULT 3
 BB546856 688 bp mRNA linear EST 26-OCT-2001
 LOCUS BB546856 RIKEN full-length enriched, 0 day neonate eyeball Mus
 DEFINITION musculus cDNA clone E130314119 3' similar to AB020886 Mus musculus
 mRNA for SSECKS, mRNA sequence.
 ACCESSION BB546856
 VERSION BB546856.2 GI:16447501
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 688)
 AUTHORS Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,
 Hiramoto,K., Hori,F., Ishi,Y., Ito,M., Kawai,J., Konno,H., Kouda,
 M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
 Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,
 D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
 Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
 Muramatsu,M. and Hayashizaki,Y.
 RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
 JOURNAL Unpublished (2001)
 COMMENT On Jul 31, 2000 this sequence version replaced gi:9618284.
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gsc.riken.go.jp,
 URL: <http://genome.gsc.riken.go.jp/>
 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
 M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
 Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura
 S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
 Hayashizaki,Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)
 Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
 Y. and Hayashizaki,Y.
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa
 K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
 Hayashizaki,Y.
 Computational Analysis of Full-length Mouse cDNAs Compared with
 Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
 Please visit our web site (<http://genome.gsc.riken.go.jp/>) for

further details.
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

FEATURES

source

```

location/Qualifiers
1..668
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="E130314119"
/clone_lib="RIKEN full-length enriched, 0 day neonate eyeball"
/tissue_type="eyeball"
/lab_host="DHI0B"
/dev_stage="0 day neonate"
/notes="Site 1: Salt; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGAGAGAGGCGCGCAACTCGAGTTTCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATTCTCGAGTTTAATTAATTAATCCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FliC I."
BASE COUNT    199 a    146 c    229 g    93 t    1 others
ORIGIN

```

Alignment Scores:

```

Pred. No.:    1.01e-71    Length:    668
Score:        1000.00    Matches:    198
Percent Similarity:    91.48%    Conservative:    6
Best Local Similarity:    88.79%    Mismatches:    17
Query Match:    12.39%    Indels:    2
DB:            10        Gaps:    2

```

US-09-902-432-4 (1-1596) x BB546856 (1-668)

```

QY 385 Gtutysmetgluaiahsgingluvalaiaagluvalahisvalserthrvagluys 404
|||
DB 3 GAGAAACGAGAGCCACCAAGAAAGTTGTCAGAGGTCCACCTGAGACCGTGAGAAAG 62
|||
QY 405 ThrGlugluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 424
|||
DB 63 ATGACCAAAAGGCGCA--GGAGAGCAGAGCTGGAAGGGGATGTGTGTGGAAGATCG 119
|||
QY 425 Gtutysmetgluaiahsgingluvalaiaagluvalahisvalserthrvagluys 444
|||
DB 120 GGGAAATCTCTGCGCCCTGAGAACTGCTGAGAACCCAGAGGTCCCGAGAACTGAG 179
|||
QY 445 ProalagluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 464
|||
DB 180 CCGTGTGAGAGAGCTGATGAAGAACCAAGAAAGTATGCTCTCGGGGTGACCAATCTCAG 239
|||
QY 465 LeuThrAspLeuSerProGluGluGluGluGluGluGluGluGluGluGluGluGlu 484
|||
DB 240 CTGACGAGATCTAAGTCTCTGAAAGAAAGATGCTACCCAAACCCCGAAGGCAATGTCAGT 299
|||
QY 485 GluValGluMetLeuSerSerGluGluGluGluGluGluGluGluGluGluGluGluGlu 504
|||
DB 300 GAGGTGAGATGCTGCTCTCTCGAGAGAGATCAAGGTACAGGGAATCCCTCGAAGAG 359
|||
QY 505 LeuPheSerSerSerGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 523
|||
DB 360 CTTTTCAGAGAGTTTGGGCTTAAAGAAAGCTCTCGGGGAAAGAAAGAGGGAAGAGAGA 419
|||
QY 524 Gtutysmetgluaiahsgingluvalaiaagluvalahisvalserthrvagluys 543
|||

```

```

DB 420 GAGGCGGGGAGATGAAGAGCCAGAGAAATACCAACATTCACCAAGGATCCCAAG 479
|||
QY 544 SerAlaaspGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 563
|||
DB 480 AGTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 539
|||
QY 564 ThrCysGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 583
|||
DB 540 GCGTGTCTGGAGAAAGGGGCGCATCGGAAGCCACCCAGAAAGCGAAGTGAAGAGAGCG 599
|||
QY 584 ThrSeraspGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 603
|||
DB 600 ACTTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATG 659
|||
QY 604 ValThrPro 606
|||
DB 660 GTGACACCC 668
|||
RESULT 4
BI652823
LOCUS
DEFINITION
603298753F1 NCI_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5339244 5',
mRNA sequence.
ACCESSION
BI652823.1 GI:15567055
VERSION
BI652823.1
KEYWORDS
SOURCE
ORGANISM
Mus musculus
house mouse.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Scuriongnathi; Muridae; Murinae; Mus.
1 (bases 1 to 661)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
JOURNAL
TITLE
Unpublished (1999)
AUTHORS
Contact: Robert Strausberg, Ph.D.
Email: cgsrbs-remail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LMNL at:
http://image.llnl.gov
Plate: LMNL1861 row: j column: 13
High quality sequence stop: 661.
FEATURES
source
1..661
location/Qualifiers
/organism="Mus musculus"
/strain="129/C57BL/6J,FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5339244"
/clone_lib="NCI CGAP Mam3"
/tissue_type="tumor_gross tissue"
/dev_stage="10 months"
/lab_host="DHI0B"
/notes="Organ: mammary; Vector: pCMV-SPORT6; Site 1: Salt;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH
Reference for transgenic model: Xu et al., Nature Genetics
22, 37-43 (1999)."
BASE COUNT    183 a    179 c    201 g    98 t
ORIGIN

```

Alignment Scores:

```

Pred. No.:    3.22e-64    Length:    661
Score:        908.00    Matches:    187
Percent Similarity:    90.91%    Conservative:    13
Best Local Similarity:    85.00%    Mismatches:    20
Query Match:    11.25%    Indels:    1
DB:            13        Gaps:    0

```

US-09-902-432-4 (1-1596) x BI652823 (1-661)

```
QY 876 ValIleAspGlyThrArgAlaValThrSerValGluGluArgSerProSerThrIleSer 895
Db 3 GTCAATTGATGGACCAAGCA-GTCAACAGTCCGCAAGAGCGGTCCCTTCGTGGATATCT 61
QY 896 AlaSerValThrGluProLeuGluHisThrAlaGlyGluAlaMetProProValGluGlu 915
Db 62 GCTTCCATGACAGAACTCTTGAGCAGCAGAGGAGTGCCACACCGCTGTGGAGAG 121
QY 916 ValThrGluLysAspIleAlaGluGluThrProValLeuThrGlnThrLeuProGlu 935
Db 122 GTCACGTAAAGACATCACTCGAGAGCAACTCTGCACTCGCCAGATTTACCAAGG 181
QY 936 GlyLysAspAlaHisAspMetValThrSerGluValAspPheThrSerGluAlaVal 955
Db 182 GGCAAGATGCCATGACGACATAGTCACCAAGTGAAGTGAATTTACCTCAGAAGCAGTG 241
QY 956 ThrAlaThrGluThrSerGluAlaLeuArgThrGluGluValThrGluAlaSerGlyAla 975
Db 242 ACAGCCGACAGAAACCAAGAGGCGCTCCGCGTGAAGAACTTACCGAAGCATCAGGGCA 301
QY 976 GluGluThrThrAspMetValSerAlaValSerGlnLeuThrAspSerProAspThrThr 995
Db 302 GAAGACACCAACATGAGTGTCTGCAGTTTCCAGCTGTCGCACTCCCGGACACCA 361
QY 996 GluGluAlaThrProValGlnGluValGluSerGlyValLeuAspThrGluGluGlu 1015
Db 362 GAGGAAGCAACCCAGTTCCAGGAGTAGAGGTTGCGATGTAGTACCGAAGAACAGGAG 421
QY 1016 ArgGlnThrGlnAlaLeuGlnAlaValAlaAspLysValLysGluGluSerGlnVal 1035
Db 422 CGCCAGACGAGCGCGCTCCCAAGCGTTGCGACAAAGTGAAGAGGACTCCCAAGTG 481
QY 1036 ProAlaThrGlnThrValGlnAlaArgThrGlySerLysAlaLeuGluValGluVal 1055
Db 482 CTGTCAACCCAGACTCTGCGAGAGAGCGGCCGGAAGACATGGAGAGGTGGAGAGGTA 541
QY 1056 GluGluAspSerGluValLeuAlaSerGluLysGluLysAspValMetProLysGlyPro 1075
Db 542 GAGGAGACTCCGAGGTGTGCTACCGAGAAAGAGAGATGTTGTGCGGAAGGACCC 601
QY 1076 ValGlnGluAlaGlyAlaGluHisLeuAlaGlnGlySerGluThrGlyGlnAlaThrPro 1095
Db 602 GTCAGGAAGCTGAACTGAGCATCTTGCACAGGCTTCCAGAGCTGTACAGGCTACCCCA 661
```

```
RESULT 5
AL045541
LOCUS AL045541 674 bp mRNA linear EST 29-FEB-2000
DEFINITION DKF2p434K035_r1 434 (synonym: htes3) Homo sapiens cDNA clone
          DKF2p434K035_5', mRNA sequence.
ACCESSION AL045541
VERSION AL045541.1 GI:5433672
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE Duesterhoeft,A., Lauber,J., Mewes,H.W., Gassenhuber,J. and Wiemann
          ,S.
          EST (Duesterhoeft, et al.)
          Unpublished (1999)
CONTACT Duesterhoeft A
MIPS
Am Klopferpitz 18a D-82152 Martinsried, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by Olagen (Hilden/Germany) within the cDNA sequencing
consortium of the German Genome Project.
No 5' sequence available.
This clone (DKF2p434K035) is available at the RZPD in Berlin.
```

Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin- Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES
source

```
1..674
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKF2p434K035"
/clone_lib="434 (synonym: htes3)"
/tissue_lib="testis"
/dev_stage="adult"
/lab_host="DH10B"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
```

BASE COUNT 213 a 139 c 212 g 109 t 1 others
ORIGIN

Alignment Scores:

Pred. No.: 5,28e-60 Length: 674
Score: 856.50 Matches: 170
Percent Similarity: 86.04% Conservative: 21
Best Local Similarity: 76.58% Mismatches: 30
Query Match: 10.61% Indels: 1
DB: 9 Gaps: 1

US-09-902-432-4 (1-1596) x AL045541 (1-674)

```
QY 424 ThrGlyGluSerLeuProProGluLysLeuAlaGluProGlnGluValProGlnGluAla 443
Db 4 ACAGCAGGCTGTGTCGCACTGAAGAAATGGTTGAAATGGATGCAGAACCTCAGGAAGCC 63
QY 444 GluProAlaGluGluLeuMetLysSerArgGluMetCysValSerGlyGlyAspHisThr 463
Db 64 GAACCTGCCAAGAGAGCTGTGTGAAGCTCAAGAAACGCTGTGTTCGAGAGGACCTTACA 123
QY 464 GlnLeuThrAspLeuSerProGluGluLysThrLeuProLysHisProGluGlyIleVal 483
Db 124 CAGGGAGCTGACCTGCTGATGAGAAAGGTGCTGTCCAAACCCCGAAGCGGTGTG 183
QY 484 SerGluValGluMetLeuSerSerGlnGluArgIleLysValGlnGlySerProLeuLys 503
Db 184 AGTGAGTGGAAATGCTGTCATCACAGGAGAGAAATCAAGGTGCAGGAAAGTCCACTAAAG 243
QY 504 LysLeuPheSerSerSerGlyLysLysLysLeuSerGlyLysGlnLysGlyLysArg 523
Db 244 AAGCTTTTACCAGCACTCGCTTAAAAAGCTTTCTGGAAGAAACAGAAAGGGAAGA 303
QY 524 GlyGlyGlyAspGluProGlyGluTyrGlnHisIleHisThrGluSerProGlu 543
Db 304 ---GGAGGAGGAGACGAGGAATCAGGGGAGCACACTCAGGTTCCAGCCGATTCCTCGGAC 360
QY 544 SerAlaAspGluGlnLysGlyGluSerSerAlaSerSerProGluGluProGluGluThr 563
Db 361 AGCCAGGAGGAGCAAAAGGGCGAGAGCTCTGCTCATCCCTCAGGAGCCCGAGGAGATC 420
QY 564 ThrCysLeuGluLysGlyProLeuGluAlaProGlnAspGlyGluAlaGluGluThr 583
Db 421 ACCTGTCTCGAAAAGGGCTTTAGCCGAGGTGCAGCAGGATGGGGAAGCTGAAGAAGGAGCT 480
QY 584 ThrSerAspGlyGluLysArgGluGlyIleThrProTrpAlaSerPheLysLysMet 603
Db 481 ACTTCGATGGAGAGAAAAAAGAGAGGTGTCACTCCCTGGGCATCATTTCAAAAAGATG 540
QY 604 ValThrProLysLysArgValArgArgProSerGluSerAspLysGluGluLeuGlu 623
Db 541 GTGAGCCCAAGAGCGTGTAGACGGCTTCGAAAGTGATATAAGAGATGAGCTGGAC 600
QY 624 LysValLysSerAlaThrLeuSerSerThrAspSerThrValSerGluMetGlnAspGlu 643
Db 601 AAGGTCAAGAGCGCTACCTGTGTTCCACCGAGAGCACAGCCTCTGAAATGCAAGAAGAA 660
QY 644 ValLys 645
Db 661 ATGAAA 666
```

RESULT 6
 BM804332 1117 bp mRNA linear EST 05-MAR-2002
 LOCUS AGECOURT 6462209 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5581233
 DEFINITION 5', mRNA sequence.
 ACCESSION BM804332
 VERSION BM804332.1 GI:19121155
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 1117)
 AUTHORS NIH-MGC http://mgi.mgi.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Place: LAM12340 row: m column: 10
 High quality sequence stop: 623.
 Location/Qualifiers
 1. 1117
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_1b="NIH_MGC_67"
 /clone_1b="NIH_MGC_67"
 /tissue_type="retinoblastoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo
 Average insert size 1.75 kb. Library constructed by Life
 Technologies."
 BASE COUNT 361 a 257 c 339 g 159 t 1 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.42e-59 Length: 1117
 Score: 854.50 Matches: 206
 Percent Similarity: 66.58% Conservative: 37
 Best Local Similarity: 56.44% Mismatches: 77
 Query Match: 10.58% Indels: 46
 DB: 14 Gaps: 10
 US-09-902-432-4 (1-1596) x BM804332 (1-1117)
 QY 291 G|TTPALG|TTPARGLVSLYSTRSERPhelYSLYSErYSGluAAspLeuGlu 310
 DB 1 GGACGGCTGGCTGGCGCAAAAGACAGTTTCAGGAGCCGAAGAGATGAGTGGAA 60
 QY 311 ThAlAGluLysArgLysGluGluGluAGluLysValAspLugluGluLysGluLys 330
 DB 61 GCTTCAGAGAGAAAAGAACAGAGCCGAAAGTATACACAGAGAAAGCGAAAG 120
 QY 331 ThrGluProAlaSerGluGlu-----GluGluProAla 341
 DB 121 GCAGAGCTTGCCTCCGAGAAATGACCGCTCCGAGAAAGCCACACGAGGCGGCA 180
 QY 342 GluAspThrAspGluAlaArgLysSerAlaAspThrGluLysValGluLeuProLeuGlu 361
 DB 181 GAAAGTCCCAACAGCCCGGTTATGAGCTGATGATGAGAAAGTGAAGTCCCTCAGAG 240
 QY 362 AspGluValGlyAspLeuGluAlaSerSerGluGluLysCysAlaProLeuAlaThrGlu 381
 DB 241 GAGCAAGTACGTGGCTCCGAGGAGCCTTCGAAAGAGAAACCTGCTCGTGGGACAGAA 300
 QY 382 ValPheAspGluLysMetGluAlaHisGln---GluValValAlaGluValHisValSer 400

DB 301 GTGTTGATGAGAAATATGAAGTCCACCAAGAGAGTTGGCCGAGTCCAGCTCAGC 360
 QY 401 ThrValGluLysThrGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 420
 DB 361 ACCGTGAGAGAGAGAACCGAAGAGCAG-----AAAACGAG 396
 QY 421 ValGluGlyThrGluGluSerLeuProGluLysLeuAlaGluProGluGluValPro 440
 DB 397 GTGAAAGAAACAGCAGGCTGTGTGCAGCTGAAAGATTGTGAATGATGACAGAACCT 456
 QY 441 GlnGluAlaGluProAlaGluGluGluMetLysSerArgLysGluMetCysValSerGly 460
 DB 457 CAGGAAGCCGAACCTGCGCAAGAGCTGTGAAGCTCAAGAAAGAGTGTTCGGAAG 516
 QY 461 AspHisThrGlnLeuThrAspLeuSerProGluGluLysThrLeuProLysHisProGlu 480
 DB 517 GACCTTACACAGGAGCTGACCTGATGATGAGAAAGTGTGTCACAAACCCCGAA 576
 QY 481 GlyLeuValSerGluValGluMetLeuSerSerGlnGluAlaGlyLeuValGlnGlySer 500
 DB 577 GCGCTTGAGAGAGTGAAGAAATGCTGTATCACAGAGAGAAATGAAGTCCAGGGAAGT 636
 QY 501 ProLeuLysLysLeuPheSerSerSerGlyLeuLysLysLysLysLysLysLysLys 520
 DB 637 CCATTAAGAAAGCTTTTACAGCAGCTGCTTAAAGAGCTTTCTGGGAAAGAACAGAG 696
 QY 521 GlyLysArgGlyGlyGlyLysAspGluGluProGlyGlyGlyGlyHisLeuHisThrGlu 540
 DB 697 GGAAGAAAGAG---GAGA-GGAGAGAGAAATCCGGGAGACACACTTCAAGCCGCA 752
 QY 540 UserProGluSerAlaAspGluGluGlyGlyGlyLysSerAlaSerProGluGlu-P 560
 DB 753 TTCTCCGAGCAGCCGAGAGAGCCGAAAGGCGAGAGCTCTGCTCATCCCTGAGAGAGC 812
 QY 560 roGluGluThrThrLysLysLysLysGlyProLeuGluAlaProGluAspGlyGluAlaG 580
 DB 813 CGAAGAAAGCCCTGCTGTGAAAGAGGCTTATC-----CAAAG---GGCACCAAG 863
 QY 580 LngLuglyThrThr-----SerAspGlyGluLysLysArgGluGly 594
 DB 864 AATGGGAAACCTGTAAAGAAAGAACTACTTTCGAGGGGGAAGAAAGAAAGAGGGG 923
 QY 594 LeThrPro---ThrAlaSerPheLysLys-----MetValThrProLysLysArgVal- 610
 DB 924 CCACCTCCCGGGGGCTTCTTCCAAAGAGGGGGAGGCCCCCAAAAGAGCGGTGT 983
 QY 611 -----ArgArgProSerGluSerAspLysG 619
 DB 984 TAAACGGCCCTTCCGAAAGGGGTATTAACAAAGAAAGAACTGGGAAACACGGGGTAC 1043
 QY 619 LngLugluLeu 622
 DB 1044 AAGAAAGCCTC 1054
 RESULT 7
 DB8456 1153 bp mRNA linear EST 01-APR-1998
 LOCUS DB8456 Mouse ICR testis (J.Tsuschida) Mus musculus cDNA clone A-12,
 DEFINITION mRNA sequence.
 ACCESSION DB8456
 VERSION DB8456.1 GI:1669580
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 1153)
 AUTHORS Takeo,M.M., Arai,Y., Matsunaga,A., Yokoi,A., Tsuschida,J., Nishina
 Y., Nozaki,M., Tanaka,H., Koga,M., Uchida,K., Matsunaga, Okuyama
 A., Rochelle,J.M., Nishimune,Y., Matsui,M. and Seidlin,M.
 TITLE Mapping of eight testis-specific genes to mouse chromosomes
 JOURNAL Genomics 46 (1), 138-142 (1997)

MEDLINE

COMMENT

98066773
Contact: Junji Tsuchida
Department of Science for Laboratory Animal Experimentation
Research Institute for Microbial Diseases, Osaka University
3-1 Yamadaoka, Suita, Osaka 565, Japan
Tel: 06-879-8338
Fax: 06-879-8339
Email: tsuchida@biken.osaka-u.ac.jp.

FEATURES

source

1. .1153
/organism="Mus musculus"
/strain="ICR"
/db_xref="taxon:10090"
/clone="A-12"
/clone_lib="Mouse ICR testis (J.Tsuchida)"
/sex="male"
/tissue_type="testis"
BASE COUNT 367 a 285-c 299 g 201 t 1 others
ORIGIN

Alignment Scores:

Pred. No.: 1.78e-59 Length: 1153
Score: 853.50 Matches: 182
Percent Similarity: 80.08% Conservative: 19
Best Local Similarity: 72.51% Mismatches: 43
Query Match: 10.57% Indels: 8
DB: 14 Gaps: 2

US-09-902-432-4 (1-1596) x D88456 (1-1153)

QY 1351 GlySerProSerLeuProAspGlnAspLysAlaGlyCysIleGluValGlnValGlnSer 1370
Db 3 GGAAGCCCTTCTCTCCAGACCAAGACAAAGCAGATTGCGATAGAGTTCAAGTTCAAGC 62
QY 1371 LeuAspThrValThrGlnThrAlaGluAlaValGluLysValIleGluThrValVal 1390
Db 63 TCAGACACACAGCTCACTCAACACACGAGCTGTGAAGAAGTCGAGAACTGTGGCA 122
QY 1391 IleSerGluThrGlyGluSerProGluCysValGlyAlaHisLeuLeuProAlaGluLys 1410
Db 123 ACTTCAGAGATGGATGAAGATTGTGGAGTGTGCAGGTGCGCAATCATTTACCAGCTGAGAAG 182
QY 1411 SerSerAlaThrGlyHisThrThrLeuGlnHisAlaGluAspThrValProLeuGly 1430
Db 183 CTCTCGAAACCGGTGCTACGGAGCTCTTCAGCATGGAGGACACCGTGCCCGCAGGNG 242
QY 1431 ProGluSerGlnAlaGluSerIleProIleValThrProAlaProGluSerThrLeu 1450
Db 243 CCTGAGTCTCAGCAGAGTCCATCCCAATATAGTAATCTCTGCTCTGAAGCATCCTA 302
QY 1451 HisProAspLeuGlnGlyGluIleSerAlaSerGlnArgGluArgSerGluGluAsp 1470
Db 303 CATCTGACCTTCAAGAGAAGTGAGCGCATCCAGAAACAGAGATCAGATGAAGATAAC 362
QY 1471 LysProAspAlaGlyProAspAlaAspGlyLysGluSerThrAlaIleGluLysValLeu 1490
Db 363 AAGCCAGATGCTGCTCTGATGCTGCGCGAAGAGAGAGTGCAGCAGAGAGAAATCCTC 422
QY 1491 LysAlaGluProGluIleLeuGluLeuSerLysSerHisLysIleValLeuAsnVal 1510
Db 423 AGGCTGAACCTGAGATCTTGGAACTTGAGAGTAAGAGCAATAGATTGTCCAGAGTGC 482
QY 1511 IleGlnThrAlaValAspGlnPheAlaArgThrGluThrAlaProGluThrHisAlaTyr 1530
Db 483 ATCCAGACAGCGCTCGACCAAGTTTGACAGTACAGAAACACCCCGAATCCAGCTTCT 542
QY 1531 AspSerGlnThrGlnValProAlaCysArgLeuAspSerArgGluProAsnArgCysTyr 1550
Db 543 GATTTCAGAAATCAGGTTCCTGTG-----ATGCA-GCCAGGAGGACACACAGATGCTGG 595
QY 1551 ThrLysMetLys-----AspAlaLysMetLysHisProValProGlnPro 1565
Db 596 ACAAGATGAAGAGGACCTTCAAGTCTCCCCCAAGATGGAACACATCAGTGCCTAGGCC 655

QY 1566 ArgGluAspLeuGlnValLeuThrValLeuGluAlaThrAlaGlnProArgLysCysLeu 1585
Db 656 AGGAGGACTTGGGTTTCTGATAGTTCTGAAGGCATGACGAGGCTTCAGAAATGATCA 715
QY 1586 ProArgLeuGlnLeuLysAlaProValSerLys 1596
Db 716 CCACGCTTGCAGTTGAAGAGTGCAGGTGTCAAAG 748
RESULT 8
AI225639 610 bp mRNA linear EST 29-OCT-1998
LOCUS ujl309.y1 Sugano mouse kidney mkia Mus musculus cDNA clone
DEFINITION IMAGE:1907968 5' similar to TR:Q62766 Q62766 SSECKS. ;, mRNA
sequence.
ACCESSION AI225639
VERSION AI225639.1 GI:3808692
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 610)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Scheilenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:976164
Possible reversed clone: similarity on wrong strand
Seq primer: custom primer used
High quality sequence stop: 460.
Location/Qualifiers
1. .610
/organism="Mus musculus"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1907968"
/clone_lib="Sugano mouse kidney mkia"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: kidney; Vector: pME18S-FL3; Site_1: DraIII (CACATGTG); Site_2: DraIII (CACATGTG); 1st strand cDNA was primed with an oligo(dT) primer
[ATGTGGCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTTCGCTACTGG], digested and cloned into distinct DraIII sites of the pME18S-FL3 vector (5' site CACATGTG, 3' site CACATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTCTGCTCTAAAGCTGG and 3' end primer CGACTGCAGCTCGAGACA."
BASE COUNT 172 a 163 c 185 g 89 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 3.7e-57 Length: 610
Score: 821.00 Matches: 167
Percent Similarity: 88.61%

RESULT 10
 AI796191/c
 LOCUS
 DEFINITION
 wh43905.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:2383544 3',
 similar to TR:000498 000498 MYASTHENIA GRAVIS AUTOANTIGEN GRAVIN ;,
 mRNA sequence.
 ACCESSION
 AI796191
 VERSION
 AI796191.1 GI:5361654
 KEYWORDS
 EST.
 SOURCE
 human.
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 796)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaps@mail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 cDNA Library Prepared by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
 Insert Length: 1103 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence stop: 458.
 Location/Qualifiers
 1. .796
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2383544"
 /clone_lib="NCI_CGAP_Kid11"
 /lab_host="DH10B"
 /notes="Organ: Kidney; Vector: pT7T3D-Pac (Pharmacia) with
 a modified polylinker; Site 1: Not I; Site 2: Eco RI;
 Plasmid DNA from the normalized library NCI_CGAP_Kid3 was
 prepared, and es circles were made in vitro. Following HAP
 purification, this DNA was used as tracer in a subtractive
 hybridization reaction. The driver was PCR-amplified cDNAs
 from a pool of 5,000 clones made from the same library
 (clones IDs 1322376-1323911, 1456007-1456775, and
 1500552-1502855). Subtraction by Bento Soares and M.
 Fatima Bonaldo".
 BASE COUNT 140 a 216 c 173 g 266 t 1 others
 ORIGIN

FEATURES

source

Alignment Scores:
 Pred. No.: 4, 04e-54 Length: 796
 Score: 785.50 Matches: 173
 Percent Similarity: 70.61% Conservative: 24
 Best Local Similarity: 62.01% Mismatches: 54
 Query Match: 9.73% Indels: 28
 DB: 9 Gaps: 6
 US-09-902-432-4 (1-1596) x AI796191 (1-796)
 QY 55 AsnGlyGlnLeuSerValAsnGlyValAlaGluGln----- 67
 Db 796 AATGGTCTGCTGCNACCATCTATGGCGTAGCTGGACAGATGAGCTCAGCGCTCAGGAGG 737
 QY 68 -----GlyAspValHisValGlnGluGlnGluGlnGln 80
 Db 736 GTAACCTAGATGTCAGGAAGAGTCCCTGTCGGTCAAGGAGTCTCTAATCAGCAT---GAA 680
 QY 81 GluGluGluValValAspGluAspValGlyGlnArgGluSerGluAspValArgGluLys 100
 Db 679 GAAGAAGAAGTCAATTGTCGGCGAGGTGTATCATATGACTCTGAAGATGTGACGACAG 620

QY 101 AspArgValGluGluMetAlaAsnSerThrAlaValGluAspIleThrLysAspGly 120
 Db 619 GACTCCGATACAGAGATGACTACTAAGTCAGCGGTAGTTTCACGACATCATGATGTTGG 560
 QY 121 GlnGluGluThrSerGluIleGluGlnIleProAlaSerGluAsnValGluGlu 140
 Db 559 CAGGAGAGACACCCGATATATCGAACAGATTCCTCTTCAGAAAGCAATTAGAGAG 500
 QY 141 MetValGlnProAlaGluSerGlnAlaAsnAspValGlyPheLysLysValPheLysPhe 160
 Db 499 CTAACAACACCCACTGAGTCCAGGCTAATGATATTGGATTTAAGAAGGTGTTTAAGTTT 440
 QY 161 ValGlyPheLysPheThrValLysLysAspLysAsnGluLysSerAspThrValGlnLeu 180
 Db 439 GTTGGCTTTAAATTCATCTGTGAAAAAGGATAAGACAGAGAGCCTGACACTGTCCAGCTA 380
 QY 181 LeuThrValLysLysAspGluGlyGluAlaGluAlaSerValGlyAlaGlyAspHis 200
 Db 379 CTCACTGTGAGAAAGATGAGGGGAGGAGCA-----GCAGGGCTGGCGACCAC 329
 QY 201 GlnGluProSerValGluThrAlaValGlyGluSerAlaSerLysGluSerGluLeuLys 220
 Db 328 CAGGACCCAGCCCTT-----GGGGCTGGAGAAGCAGCATCCAAAGAAAGCGAACCCAAA 275
 QY 221 GlnSerThrGluLysGlnGluGlyThrLeuLysGlnGluGlnSerSerThrGluIlePro 240
 Db 274 CAATCTACAGAGAAACCCGAAAGAGACCCCTGAAGCGTGAGCAAAAGCCACGAGAAATTTCT 215
 QY 241 -LeuGlnAlaGluSerAspGlnAlaAlaGluGluGluAlaLysAspGluGluGluLys 260
 Db 214 CCCCCAAGCGGAATCTGGCCAAAGCAGTG---GAGGAATCAAGAGAGAGAGAGAGAA 158
 QY 260 sGlnGluLysGluProThrLysSerProGluSerProGluSerProValAsnSerGluTh 280
 Db 157 ACAAGAAAGAAACCTAGCAGTCTCCAGATCTCCGACTAGTTCCGTGACCACTGAAC 98
 QY 280 rThrSerSerPheLysLysPhePheThrHisGlyTrpAlaGlyTrpArgLysLysThrSe 300
 Db 97 AGGATCAACCTTCAAAAATTTCTCACTCAAGGTTGGCGCGCTGGCGCAAAAAGGCCAG 38
 QY 300 rPheLysLysSerLysGluAspLeuGluThrAlaGluLysArgLysGluGln 318
 Db 37 TTTCAGGAAGCCGAGGAG-----GAAAGGAGGAGGAGCTTCAG 1
 RESULT 11
 BO717711
 LOCUS BO717711 873 bp mRNA linear EST 16-JUL-2002
 DEFINITION AGENCOURT_8234483 Lupski_sympathetic_trunk Homo sapiens cDNA clone
 IMAGE:6188718 5', mRNA sequence.
 ACCESSION BO717711
 VERSION BO717711
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 873)
 NIH-MGC <http://mgc.nci.nih.gov/>
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaps@mail.nih.gov
 Tissue Procurement: Dr. James R. Lupski
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM13584 row: m column: 07
 High quality sequence stop: 627.
 Location/Qualifiers
 1. .873
 FEATURES
 source

QY 743 ValSerThrTrpGluSerPheLysArgLeuValThrProArgLysLysSerLysSerLys 762
 Db 123 GTTTCACCTGGGAGTCATTAAAGGTTAGTCACGCCAGAAAAAATCAAAAGTCCAAG 182
 QY 763 LeuGluLysAlaGluAsp-----SerSerValGluGlnLeuSerThrGlu 778
 Db 183 CTGGAGAGAAAGCGAGACTCCATAGCTGGGTCTGGTGTAGAACATTCCTCACTCCAGAC 242
 QY 779 IleGluProSerArgGluSerTrpValSerIleLysLysPheLysProGlyArg 798
 Db 243 ACTGAACCCGTTAAAGAAATCTCGGTCTCAATCAAGAAGTTTATCTCTGGACGAAG 302
 QY 799 LysLysArgAlaAspGlyLysGlnGlnAlaThrValGluAspSerGlyProValGlu 818
 Db 303 AGAAAGAGCCAGATGGGAACCAAGAACGCCCTCTTGAAGACGCGAGGCCCAACAGGG 362
 QY 819 IleAsnGluAspAspProAsnValProAlaValProLysLysThrGluTyrAsnAlaVal 838
 Db 363 GCCACGAGATGACTCTGATGTCCTCGGCGGTGGTCTCTCTGTGATGATGATGCTGTA 422
 QY 839 GluArgGluLysMetGlu-----AlaGlnGlyAsnThrGluLeuProGlnLeu 855
 Db 423 GAAAGGAGAAATGGAGGCACAGCAAGCCCAAAAGCGCAGAGCAGCCGACGACAAG 482
 QY 856 GlyAlaValThrValSerGluLysSerLysThrLeuValHisThrValSerValala 875
 Db 483 GCAGCCACTGAGGTGTCCTCAAGGAGCTCAGCAGAGTCAGGTTCATATGATGGCAGCAGCT 542
 QY 876 ValIleAspGlyThrArgAlaValThrSerValGluGluArgSerProSerTrpIleSer 895
 Db 543 GTCTGCTACGGGACGAGGAGCTACCATATTGAAAGAGGTCTCTCTCTTGATATCT 602
 QY 896 AlaSerValThrGluProLeuGluHisThrAlaGlyGluAlaMetProProValGluGlu 915
 Db 603 GCTTCAGTGACAGAACCTCTTGACCAAGTAGAGCTGAAGCCGACCTGTTAACTGAGGAG 662
 QY 916 ValThrGluLysAspIleIleAlaGluGluThrPro 927
 Db 663 GTATTGAAAGAGAGTAATTGCAGAGAAGAACCC 698

RESULT 13
 AII156819 530 bp mRNA linear EST 30-SEP-1998
 LOCUS ui44c01.y1 Sugano mouse embryo mewa Mus musculus cDNA clone
 DEFINITION IMAGE:1885248 5' similar to TR:Q62766 Q62766 SSECKS. ;, mRNA
 sequence.
 ACCESSION AII156819.1 GI:3685288
 VERSION EST.
 KEYWORDS house mouse.
 SOURCE Mus musculus
 ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 530)
 AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
 Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
 Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
 Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
 Waterston, R.
 TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:969572
 Possible reversed clone: similarity on wrong strand

FEATURES

Seq primer: custom primer used
 High quality sequence stop: 447.
 Location/Qualifiers
 1..530
 /organism="Mus musculus"
 /strain="C57BL"
 /db_xref="taxon:10090"
 /clone="IMAGE:1885248"
 /dev_stage="embryo, 14 dpc"
 /lab_host="DH10B"
 /note="Vector: pME18S-FL3; Site 1: DraIII (CACTGNGTG);
 Site 2: DraIII (CACCATGTG); 1st strand cDNA was primed
 with an oligo(dT) primer [ATGCGCCCTTTTITTTTTTTTTT];
 double-stranded cDNA was ligated to a DraIII adaptor
 [TGTGGCTCTACTGG], digested and cloned into distinct DraIII
 sites of the pME18S-FL3 vector (5' site CACTGNGTG, 3' site
 CACCATGTG). XhoI should be used to isolate the cDNA
 insert. Size selection was performed to exclude fragments
 <1.5Kb. Library constructed by Dr. Sumio Sugano
 (University of Tokyo Institute of Medical Science).
 Custom primers for sequencing: 5' end primer
 CTCTGCTCTAAAAGCTGGC and 3' end primer
 CGACCTGCAGCTCGAGACA."
 BASE COUNT 151 a 141 c 164 g 74 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,84e-51 Length: 530
 Score: 748.00 Matches: 152
 Percent Similarity: 92.61% Conservatives: 11
 Best Local Similarity: 86.36% Mismatches: 13
 Query Match: 9.27% Indels: 0
 DB: 9 Gaps: 0
 US-09-902-432-4 (1-1596) x AII156819 (1-530)

QY 950 PheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeuArgThrGluGluVal 969
 Db 3 TTTTACCTCAGAACGAGTCACAGCGCAGAAACCCACAGAGCGCTCCGCGCTGAAGAATC 62
 QY 970 ThrGluAlaSerGlyAlaGluThrThrAspMetValSerAlaValSerGlnLeuThr 989
 Db 63 ACCGAAGCATCAGGGGACAGAGAGACACACATGCTGTCTCAGTTTCCCCAGCTGTC 122
 QY 990 AspSerProAspThrThrGluGluAlaThrProValGlnGluValGluSerGlyValLeu 1009
 Db 123 GATCTCCCGGACACACAGAGGAGGCCACCCAGTTCAGAGGTAGAGGGTGGCATGCTA 182
 QY 1010 AspThrGluGluGluArgGlnThrGlnAlaIleLeuGlnAlaValAlaAspLysVal 1029
 Db 183 GATACGGAAGAACAGGAGCGCCAGAGCGCGCTCTCCAAAGCCGTTGCAGACAAAGTG 242
 QY 1030 LysGluLysSerGlnValProAlaThrGlnThrValGlnArgThrGlySerLysAlaLeu 1049
 Db 243 AAGAGAGACTCCCAAGTGGCTTCCAAACCCAGACTCTCGACAGACAGAGGCGCCGAAGACTG 302
 QY 1050 GluLysValGluGluValGluGluAspSerGluValLeuAlaSerGluLysGluLysAsp 1069
 Db 303 GAGAGGTGGAGAGGTAGAGGAGGACTCCAGGTGCTGGCTACCGAAGAGAGAGGAT 362
 QY 1070 ValMetProLysGlyProValGlnGluAlaGlyAlaGluHisLeuAlaGlnGlySerGlu 1089
 Db 363 GTTGTGCGGAGGAGCCCGCTCGAGGAAGCTGAAACTGAGCATCTTGACAGGGCTCCGAG 422
 QY 1090 ThrGlyGlnAlaThrProGluSerLeuGluValProGluValThrAlaAspValAspHis 1109
 Db 423 ACTGTACAGGCTACCCACAGAGACCTTGAAGTTCTGAAGTTCACAGAGATGTAGACCGT 482
 QY 1110 ValAlaThrCysGlnValIleLysLeuGlnLeuMetGluGlnAla 1125
 Db 483 GCCACCATCATGCCAGGTTATCAAGCACCCAGCAGCTGATGGAACAGGCT 530

RESULT 14
BQ43215 679 bp mRNA linear EST 29-MAY-2002
LOCUS UI-M-EV0-bxg-o-07-0-UI.r1 NIH_BMAP_EV0 Mus musculus cDNA clone
DEFINITION IMAGE:5707998 5', mRNA sequence.
BQ43215
ACCESSION BQ43215.1 GI:21246327
VERSION EST.
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 679)
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
cgabs-remail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LMNC at:
<http://image.llnl.gov>
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: PYX-5.
FEATURES
source
location/Qualifiers
1. 679
/organism="Mus musculus"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:5707998"
/clone_1ib="NIH BMAP EV0"
/tissue_type="Whole Brain"
/dev_stage="embryo 15.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/note="Organ: brain; Vector: pYX-Aac; Site 1: EcoR I;
Site 2: Not I; The library was constructed according to
Bonaldi, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured mRNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was size selected according to mRNA size fraction,
ligated with EcoR I adaptor, digested with Not I, and then
cloned directionally into pYX-Aac vector. The library tag
sequence located between the Not I site and the polyA tail
, is GTGCGTGGAA. This library was created for the
University of Iowa Mouse Brain Molecular Anatomy Project
(BMAP): 'Gene Discovery in the Developing Mouse Nervous
System', supported by National Institutes of Mental Health
(NIMH), Hemin Chin, Ph.D., program coordinator."

BASE COUNT 218 a 151 c 202 g 108 t
ORIGIN

Alignment Scores:
Pred. No.: 8.14e-51 Length: 679
Score: 744.00 Matches: 164
Percent Similarity: 76.65% Conservative: 10
Best Local Similarity: 72.25% Mismatches: 47
Query Match: 9.22% Indels: 7
DB: 14 Gaps: 2

US-09-902-432-4 (1-1596) x BQ43215 (1-679)

QY 1212 Gtuprthnglglglnleuthralaalaalavalprovalleualalystrhgluval 1231
Db 2 GAACCCACAC---CAAGCGCTGGCTGCGGGGCGAGTCCCTATTCTGGCAAGGCTGAGGTG 58

QY 1232 GtGtnglglglnlgluValasprtleuaspGtlygluVallygslu----- 1247
Db 59 GGTCAAGAGGGGTGAGGCTGGCGAGTTTGTATGAGAAAGTCAAAAGACGACAGCTGTGTT 118

QY 1248 ---GtGtnglgluValpnevalHisSerGlyProAsnSerGlnlyValaAlaAspValThr 1266
Db 119 AAAGAACTGGAGGTGCTGTGCACACTGGACCCAAACACTCAAAAGACCTGACTTGACA 178

QY 1267 TyraSpSerGlnValMetGlyValAlaGlyCysGlnGlnlyGlySerThrGluValGln 1286
Db 179 CGTGACAGTGAAGTATGAAAGTGGCCAGATGTCCAGAACTGACAGATTAATGAACACAG 238

QY 1287 SerLeuSerLeuGlnGlnlyGluMetGlnThrAspValGlnlyGlyValGlySerGlnThr 1306
Db 229 AGATTATACCCCGAGAAAGAGAGATGGAAACGACGTTGAAGAGAGAGAAACAGAGACC 298

QY 1307 LysPProGlnGlnValSerGlnGlnlyGlnGlnGlnThrAlaAlaProGlnHisGlnly 1326
Db 229 AAGACAGAGCAAGCCAGTGAAGAAACATAGAGAAAGAGTGTCTCCAGACATTAAGGA 358

QY 1327 ThrTyGlyLysProValLeuThrLeuAspMetProSerSerGlnuArgGlyValAlaLeu 1346
Db 359 ACCACCTTAAGCCAGTCTTACAGCTGACATGCTCTCACTCAGAGAGGGAGAAAGGCACTG 418

QY 1347 GlySerLeuGlyGlySerProSerLeuProAspGlnAspLysAlaGlyCysIleGluVal 1366
Db 419 GGCAGCCTTGAAGAGAGAGCCCTTCTCTCCAGACCAAGCAAGAGATTGATAGAGTT 478

QY 1367 GlnValGlnSerLeuAspThrThyValThrGlnThrAlaGluAlaValGlnlyValIle 1386
Db 479 CAAGTTCAAGAGCTCAGACACACCACTCACTCAACACGAGCTGTGAAAAAGGTCCAA 538

QY 1387 GlnThrValIleSerGlnThrGlyGlySerProGlnCysValGlyValHisLeuLeu 1406
Db 539 GAACCTTGCGCACTTCAGAGATGATTAAGTTGGAGTGGCGAGTGGCAATCATTA 598

QY 1407 ProAlaGlnlySerSerAlaThrGlyGlyHisStrThrLeuGlnHisAlaGlnlyThr 1426
Db 599 CCAGCTGAGAAAGCTCTCCGAAACCGGTGGCTAGCGGACTCTTCAGCAAGAGAGA-ACC 657

QY 1427 ValPProLeuGlyProGlnSer 1433
Db 658 GTGCCCGAGGGGCTGAGTCT 678

RESULT 15
BFA00579/c 573 bp mRNA linear EST 28-NOV-2000
LOCUS UI-R-CA0-bhd-a-09-0-UI.s1 UI-R-CA0 Rattus norvegicus cDNA clone
DEFINITION UI-R-CA0-bhd-a-09-0-UI 3', mRNA sequence.
BFA00579
ACCESSION BFA00579.1 GI:11388554
VERSION EST.
KEYWORDS Norway rat.
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 573)
AUTHORS Bonaldi,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
COMMENT Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
The sequence contained an oligo-dT track that was pres
oligonucleotide that was used to prime the synthesis of
strand cDNA and therefore this may represent a bonafide
tail. cDNA Library Preparation: M.B. Soares Lab Clone d
clones will be available through Research Genetics (www
Seq primer: M13 Forward

POLYA=Yes.

FEATURES Location/Qualifiers

source 1. .573
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-CA0-bhd-a-09-0-UI"
/clone_lib="UI-R-CA0"
/lab_host="DH10B (Life Technologies)"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-CA0
library is a subtracted library derived from the following
tissues: thalamus, cerebellum, hypothalamus, medulla, pons
, midbrain, cerebral cortex, corpus striatum, testis, and
hippocampus. For a detailed description of the library
from which this clone was derived, please visit our web
site at ratest.eng.uiowa.edu. The subtraction has been
previously described in (Bonaldo, Lennon and Soares,
Genome Research 6:791-806, 1996)
TAG_SEQ=None found"

BASE COUNT 91 a 148 c 124 g 210 t
ORIGIN

Alignment Scores:

Pred. No.:	2,25e-48	Length:	573
Score:	713.00	Matches:	143
Percent Similarity:	84.54%	Conservative:	21
Best Local Similarity:	73.71%	Mismatches:	24
Query Match:	8.83%	Indels:	6
DB:	12	Gaps:	3

US-09-902-432-4 (1-1596) x BF400579 (1-573)

QY	125	SerGluIleIleGluGlnIleProAlaSerGluAsnValGluGluMetValGlnPro	144
DB	564	GCCTGAATATCGACAGATTCCTCTTCAGAAAGCAATTAGAGAGCTAACACACCC	505
QY	145	AlaGluSerGlnAlaAsnAspValGlyPheLysLysValPheLysPheValGlyPheLys	164
DB	504	ACTGAGTCCCGAGGCTAATGATTTGGATTTAAGAAGGTGTTAAGTTTGTGGCTTTAAA	445
QY	165	PheThrValLysLysAspLysAsnGluLysSerAspThrValGlnLeuLeuThrValLys	184
DB	444	TTCACTGTGAAAAGGATAAGACAGACAGCCTGACACTGTCCAGCTACTCCTGTGAAG	385
QY	185	LysAspGluGlyGluGlyAlaGluAlaSerValGlyAlaGlyAspHisGlnGluProSer	204
DB	384	AAAGATGAAGGGAGGGAGCA-----GCAGGGGCTGGCGACCCAGGACCCCGC	334
QY	205	ValGluThrAlaValGlyGluSerAlaSerLysGluSerGluLeuLysGlnSerThrGlu	224
DB	333	CTT-----GGGGCTGGAGAAGCAGCATCCAAAGAAAGCGAACCCAAACAATCTACAGAG	280
QY	225	LysGlnGluGlyThrLeuLysGlnGlnSerSerThrGluIleProLeuGlnAlaGlu	244
DB	279	AAACCCGAGAGACCCCTGAAAGCGTGAAGCAAGCCACGACCAAGCAATTTCTCCCCAGCCGAA	220
QY	245	SerAspGlnAlaAlaGluGluAlaLysAspGluGlyGluGluLysGlnGluLysGlu	264
DB	219	TCTGGCCAGCAGTG---GAGGAATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAA	163
QY	265	ProThrLysSerProGluSerProSerSerProValAsnSerGluThrThrSerSerPhe	284
DB	162	CCTAGCAAGTCTGCAGAAATCTCCGACTAGTCCCGTACCAGTGAACAGGATCAACCTTC	103
QY	285	LysLysPhePheThrHisGlyTyrAlaGlyTyrPargLysLysThrSerPheLysSer	304
DB	102	AAAAAATCTTCACTCAAGGTGGCGCGCTGGCGCAAAAAAGACCCAGTTTCAGGAAGCCG	43
QY	305	LysGluAspAspLeuGluThrAlaGluLysArgLysGluGln	318
DB	42	AAGAGGATGAAGTGAAGCTTCAGAGAAAAAAGAAAAA 1	

Search completed: December 13, 2002, 04:10:45
Job time : 4157 secs

THIS PAGE BLANK (USPTO)